



Assistance visuelle des malvoyants par traitement d'images adaptatif

Benoit Froissard

► To cite this version:

Benoit Froissard. Assistance visuelle des malvoyants par traitement d'images adaptatif. Traitement des images [eess.IV]. Université Jean Monnet - Saint-Etienne, 2014. Français. NNT : 2014STET4004 . tel-01160158

HAL Id: tel-01160158

<https://theses.hal.science/tel-01160158>

Submitted on 4 Jun 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Saint-Étienne

ED SIS Science Ingénierie Santé

Faculté de sciences

Laboratoire Hubert Curien

Assistance visuelle des malvoyants par traitement d'images adaptatif

Benoit FROISSARD

Doctorat dans la spécialité : image, vision

Directeur de thèse : Alain Trémeau, Co-encadrants : Hubert Konik, Éric Dinet

Soutenue le 27 février 2014

Membres du jury :

Evelyne Lutton – Rapporteur

Serge Miguët – Rapporteur

Frédérique Robert-Inacio – Examineur

Alain Trémeau – Directeur de thèse

Hubert Konik – Co-encadrant

Éric Dinet – Co-encadrant

Zile Liu – Responsable entreprise

Résumé

Cette thèse s'inscrit dans le cadre de l'analyse des besoins de personnes malvoyantes dans le but de développer de nouvelles solutions d'assistance visuelle en utilisant le principe de réalité augmentée. Le principal objectif de ces travaux était le développement d'un dispositif d'assistance visuelle pour faciliter la mobilité de personnes malvoyantes grâce à un algorithme qui en temps réel adapte sa réponse en fonction des besoins de l'utilisateur et de la situation dans laquelle il évolue. Le premier postulat de cette thèse est d'utiliser l'acuité résiduelle fonctionnelle des malvoyants afin de leur permettre de percevoir un certain nombre d'informations. Le second postulat est d'utiliser les seules informations pertinentes, utiles à un malvoyant, afin d'améliorer sa perception du monde dans lequel il évolue.

Le premier axe de recherche a consisté à analyser les besoins des malvoyants afin de recenser, les indices visuels utiles à rehausser afin d'améliorer la perception des malvoyants. Cette étude a été réalisée en immersion et une enquête précise a été menée auprès de malvoyants. Cette étude permet de mieux comprendre les besoins des malvoyants, les méthodes qu'ils utilisent ainsi que leurs usages des dispositifs actuels.

Le deuxième axe de recherche a consisté à concevoir et à mettre en œuvre un algorithme de traitement d'images répondant aux besoins identifiés. Un outil interactif a été développé qui permet de sélectionner la quantité de contours en fonction de la situation. Ce détecteur de contour optimise sa détection afin d'obtenir des contours plus exploitables pour les malvoyants, comparés aux résultats des autres méthodes de l'état de l'art. Cet outil a été testé et validé auprès d'une population donnée de personnes malvoyantes. Ces tests réalisés sur écran ont démontré que les contours constituent des indices visuels très importants qui peuvent être d'une grande aide dans l'analyse d'une scène.

Le troisième axe de travail a consisté au développement d'un système de réalité augmentée permettant d'utiliser ce détecteur de contours dans un contexte de mobilité. L'analyse des besoins des malvoyants vis-à-vis d'un tel dispositif nous a permis de définir les caractéristiques indispensables pour le système de calcul, la caméra et le dispositif d'affichage. Concernant le système d'affichage, l'analyse des besoins des malvoyants en termes de profil, de couleur, de largeur des contours, nous a permis de définir les paramètres nécessaires pour un affichage qui répond à ces besoins. Il a également été pris en compte dans les développements les paramètres physiologiques de chaque observateur afin que les images affichées soient visuellement confortables.

Abstract

This thesis was done in the context of visual impaired needs analysis with the goal of new augmented reality visual assistance device development. The main goal of this work was the development of assistive device for mobility of visual impaired people by the use real-time image processing algorithm which take into account user needs and the context.

The first postulate is the use of residual visual acuity of visual impaired people to allow them to see more information. The second postulate is the use only most useful information for visual impaired people for the understanding of the situation where are they.

The first research axis was an analysis of needs of visual impaired to identify visual cues, which one we need to enhance to help them. To do this, I have done an immersive stage into a re-adaptation clinic for visually impaired. I have also done a questionnaire to identify and understand how current products answers partially to their needs.

The second research axis was the development of a new image processing to answer identified needs. I have developed an interactive tools that allow the user to select the amount of edges for each situation, context. This detector optimizes its result to provide better results than others state of the art evaluated methods. This tool was evaluated and validated during visual test by visual impaired people. We have demonstrated that edges are one of the most important visual cues for visual impaired people to analyze a scene.

The third research axis was the development of the augmented reality system to use my tool in a context of mobility. To do this, I have analyzed the needs of visual impaired people for this type of devices. I have also defined main requirements of the desired device for the processing unit, the camera and the display module. For display module, I have analyzed physiologic needs of visually impaired, in terms of profile, colors, and thickness to answer their needs. I have also developed an image generation model which takes into account many physiologic parameters of each observer to create visually comfortable images.

Table des matières :

RESUME	2
ABSTRACT	3
TABLE DES MATIERES :	4
TABLE DES ILLUSTRATIONS	8
REMERCIEMENTS.....	13
INTRODUCTION	14
SYNTHESE	17
1. LE SYSTEME VISUEL ET LA BASSE VISION	18
1.1. L'œil	18
1.1.1. La formation de l'image	18
1.1.2. Critères d'analyse de la vision	20
1.1.2.1. L'acuité visuelle	20
1.1.2.2. Le champ visuel	21
1.1.2.3. La sensibilité au contraste	22
1.1.3. Amétropies.....	23
1.1.3.1. La myopie	24
1.1.3.2. L'hypermétropie	24
1.1.3.3. L'astigmatisme.....	25
1.1.3.4. La presbytie	26
1.1.4. La rétine	26
1.1.4.1. Rôle et constitution	26
1.1.4.2. Champ récepteur.....	28
1.1.5. Atteintes de la rétine	29
1.1.5.1. La dégénérescence maculaire liée à l'âge	30
1.1.5.2. Le glaucome.....	30
1.1.5.3. La rétinopathie diabétique	31
1.1.5.4. La rétinite pigmentaire	31
1.2. Le système visuel.....	32
1.2.1. Introduction	32
1.2.2. Vision monoculaire	33
1.2.3. Vision binoculaire.....	33
2. ANALYSE DES BESOINS: ASPECTS QUALITATIFS ET QUANTITATIFS.....	37
2.1. Introduction	37
2.2. Analyse quantitative des besoins	39

2.2.1.	Les enquêtes existantes	39
2.2.2.	Les produits du marché.....	43
2.2.2.1.	Les loupes optiques	44
2.2.2.2.	Les loupes électroniques	44
2.2.2.3.	Les machines à lire	45
2.2.2.4.	Les télé-agrandisseurs	46
2.2.2.5.	Les solutions informatiques.....	46
2.2.3.	Notre étude menée à partir d'un questionnaire	47
2.2.3.1.	Catégories des malvoyants interviewés	48
2.2.3.2.	Besoins exprimés par ces malvoyants	51
2.2.3.3.	Aspects ergonomiques	55
2.2.4.	Conclusion.....	57
2.3.	Immersion	58
2.3.1.	Contexte de l'immersion	58
2.3.1.1.	Présentation de l'Institut ARAMAV	59
2.3.1.2.	L'orthoptie	61
2.3.1.3.	L'ergothérapie	62
2.3.1.4.	La psychomotricité et la locomotion	63
2.3.2.	Analyse.....	66
3.	IMAGE ET CONTOURS POUR MALVOYANTS.....	69
3.1.	Etat de l'art.....	69
3.1.1.	Introduction	69
3.1.2.	Oculométrie	70
3.1.3.	Modélisation des mécanismes de la vision et assistance visuelle	72
3.1.4.	Modèles de saillance et d'apparence visuelle	73
3.1.5.	Rehaussement et anamorphose d'image	75
3.1.6.	Détection de contours	77
3.1.7.	Dispositifs d'affichage sur tête.....	78
3.1.8.	Orthèses visuelles et tactiles.....	80
3.2.	Détecteur de contours adaptatif	81
3.2.1.	Banque d'images.....	81
3.2.1.1.	Banque d'images « contours »	81
3.2.1.2.	Notre banque de test	83
3.2.2.	Sobel	85
3.2.3.	Canny	87
3.2.4.	Canny couleur	90
3.2.5.	L'aspect multi-échelle : les pyramides	91
3.2.6.	Filtrage de contours par pyramide.....	93
3.2.6.1.	Filtrage ajustable	94
3.2.6.2.	Détection de contours multi-échelle.....	95
3.2.6.3.	Fusion, sélection et restitution des contours	97
3.2.6.4.	Présentation de résultats	100
3.2.6.5.	Schéma général de l'algorithme.....	102
3.2.6.6.	Comparaisons quantitatives aux méthodes de l'état de l'art	103
3.3.	Conclusion.....	104
4.	ÉVALUATION DE L'INTERET DU REHAUSSEMENT DES CONTOURS POUR LES MALVOYANTS.....	106
4.1.	État de l'art.....	106

4.2.	Présentation du protocole expérimental	107
4.2.1.	Evaluation des contours.....	108
4.2.2.	Evaluation de la restitution	111
4.3.	Malvoyants ayant participé à l'évaluation	112
4.4.	Résultats de l'évaluation	115
4.4.1.	Introduction	115
4.4.2.	Le rehaussement des contours peut-il être une méthode d'assistance visuelle ?	116
4.4.3.	Existe-t-il un lien entre l'épaisseur des traits à utiliser pour le rehaussement des contours et l'acuité visuelle ?	117
4.4.4.	Quelles sont les méthodes d'extraction de contours les plus pertinentes pour l'assistance des malvoyants ?	120
4.4.5.	Existe-t-il des couleurs préférentielles pour rehausser les contours ?	122
4.4.6.	Existe-t-il des profils de traits pour rehausser les contours ?	125
4.4.7.	Existe-t-il des épaisseurs de traits à privilégier pour rehausser les contours ?	125
4.5.	Conclusion de l'évaluation.....	127
5.	DISPOSITIF D'ASSISTANCE.....	128
5.1.	Dispositif de restitution visuelle	128
5.1.1.	Réalité virtuelle vs Réalité augmentée	128
5.1.2.	Dispositifs à verre transparent.....	131
5.1.2.1.	Systèmes monoculaires	131
5.1.2.2.	Monoculaire vs binoculaire	136
5.1.2.3.	Modèle de génération d'image binoculaire	138
5.1.3.	Recalage des informations	141
5.1.4.	Choix du dispositif d'affichage	143
5.2.	Système d'acquisition : caméra	144
5.2.1.	Introduction	144
5.2.2.	Caméra CCD	145
5.2.3.	Caméra CMOS	146
5.2.4.	Critères de sélection	148
5.3.	Unité de calcul.....	150
5.3.1.	Implémentation logicielle	151
5.3.2.	Comparaison des systèmes.....	154
5.4.	Conclusion.....	157
	SYNTHESE	158
	CONCLUSION.....	160
	ANNEXE 1 : QUESTIONNAIRE	166
	ANNEXE 2 : LA BANQUE D'IMAGES.....	172
	ANNEXE 3 : IMAGES ET VERITE TERRAIN.....	174
	ANNEXE 4 : ANALYSE QUANTITATIVE	175

BIBLIOGRAPHIE	183
----------------------------	------------

Table des illustrations

Figure 1.1 : Représentation de l'œil	19
Figure 1.2 : Échelle logarithmique log MAR.....	21
Figure 1.3 : Acuité en fonction du champ visuel	22
Figure 1.4 : Courbe de sensibilité aux contrastes.....	23
Figure 1.5 : Schéma d'un œil myope non corrigé.....	24
Figure 1.6 : Schéma d'un œil hypermétrope non corrigé.....	25
Figure 1.7 : Organisation schématique de la rétine.....	26
Figure 1.8 : Distribution spatiale des cônes et des bâtonnets au sein de la rétine.....	27
Figure 1.9 : Sensibilité spectrale des cônes et bâtonnets.....	27
Figure 1.10 : Fonctionnement d'un champ récepteur	28
Figure 1.11 : (à gauche) Image d'une vision normale ; (à droite) d'une vision atteinte de DMLA	30
Figure 1.12 : (à gauche) Image d'une vision normale; (à droite) d'une vision atteinte de glaucome	31
Figure 1.13 : (à gauche) Image d'une vision normale ; (à droite) d'une vision atteinte de rétinopathie diabétique	31
Figure 1.14: (à gauche) Image d'une vision normale ; (à droite) d'une vision atteinte de rétinite pigmentaire	32
Figure 1.15 : Relation accommodation A (en dioptries)-convergence (C en angle métrique) du système visuel	34
Figure 1.16 : Schéma de l'horoptère expérimental et œil cyclope (à droite) ; montrant les appariements de points entre les deux yeux.	36
Figure 2.1 Modes, de lecture, utilisés par les malvoyants et les solutions techniques, données compilées par Laster à partir de l'étude handicapzéro	40
Figure 2.2 Mode d'accès à des contenus multimédia pour les personnes malvoyantes, en pourcentages, données compilées par Laster à partir de l'étude HandicapZéro	41
Figure 2.3 Besoins, en pourcents, des personnes malvoyantes par tranches d'âge, données du Ministère des affaires sociales et de la santé & INSEE - études et résultats juillet 2005 n° 416	42
Figure 2.4 Tableau comparatif des loupes électroniques, données issues des fiches techniques et catalogues des distributeurs en France,,,,,	45
Figure 2.5 Tableau comparatif des machines à lire, données issues des fiches techniques et catalogues des distributeurs en France.	46
Figure 2.6 Tableau comparatif des télé-agrandisseurs, données issues des fiches techniques et catalogues des distributeurs France.	46
Figure 2.7 Tableau comparatif des logiciels informatiques, données issues des fiches techniques et catalogues des distributeurs en France	47
Figure 2.8 Répartition de l'âge des malvoyants ayant répondu au questionnaire.....	48
Figure 2.9 Catégorie socio-professionnelles des malvoyants en pourcentage par rapport à chaque groupe.	49
Figure 2.10 Répartition des acuités monoculaires en vision de loin.....	50
Figure 2.11 Répartition des malvoyants en fonction de l'atteinte du champ visuel.	50
Figure 2.12 Taux de réponses pour chaque type d'aide visuelle et l'utilité de ces aides vis-à-vis de certains usages	51
Figure 2.13 Taux d'intention d'acquisition d'une aide visuelle.....	52

Figure 2.14 Impacts dus à la déficience visuelle.....	53
Figure 2.15 Taux de malvoyants estimant pouvoir réaliser certaines actions en autonomie... ..	54
Figure 2.16 Fonctionnalités demandées	55
Figure 2.17 Autonomie acceptable pour une aide visuelle	56
Figure 2.18 Poids du boîtier acceptable par les malvoyants pour une aide visuelle.	56
Figure 2.19 Poids acceptable pour une paire de lunettes d'aide visuelle.	56
Figure 2.20 : Organisation de l'ARAMAV.....	60
Figure 2.21 : Exemple de passage piéton avec repère podotactile (à gauche) et une place comportant différentes textures de sol (à droite).....	65
Figure 2.22 : Exemple de trottoir large (à gauche) et de trottoir étroit (à droite).	65
Figure 2.23 : Zone explorable avec une canne blanche.	65
Figure 3.1 : Exemple de suivi de position du regard.....	70
Figure 3.2 : Exemple de dispositifs d'oculométrie : (à gauche) un dispositif pour la recherche ; (à droite) un dispositif pour tablette et PC.	71
Figure 3.3 : Exemple de fixations et saccades lors de la lecture.....	72
Figure 3.4 : (à gauche) Image d'origine ; (à droite) reconstruction de l'image en conservant uniquement les zones saillantes de l'image [48, 47].....	75
Figure 3.5 : Représentation des différentes fréquences d'une matrice DCT	76
Figure 3.6 : Rehaussement MPEG : (à gauche) l'image rehaussée ; (à droite) l'image d'origine en symétrie.....	77
Figure 3.7 : Exemple de HMD, pour la restitution de contours, développé par l'équipe d'Eli Peli.....	78
Figure 3.8 : De gauche à droite: une image de l'environnement, la représentation tactile de l'environnement, et l'interface tactile (TactiPad, ISIR, Université Paris 6).	80
Figure 3.9 : (à gauche) Une scène réelle ; (à droite) la restitution de cette scène par le stimulateur rétinien.....	81
Figure 3.10 : Images de BSDS300, de gauche à droite : l'image originale en couleur, la carte de contours d'un observateur, la carte de contours d'un deuxième observateur, et la fusion des cartes de contours des observateurs	82
Figure 3.11 : Images de BSDS300, de gauche à droite et de haut en bas: l'image originale en couleur, la fusion des cartes de contours des observateurs, la carte de contours d'un observateur et la carte de contours d'un deuxième observateur.....	83
Figure 3.12 : Images de BSDS300, de gauche à droite : l'image originale en couleur, la carte de contours d'un observateur, la carte de contours d'un deuxième observateur, et la fusion des cartes de contours des observateurs	83
Figure 3.13 : Quelques images de la banque, la banque d'image complète est en annexe 2 ...	85
Figure 3.14 : (à gauche) Image originale ; (à droite) image filtrée par le filtre de Sobel et seuillée (valeur 100).	87
Figure 3.15 : (à gauche) Image originale ; (à droite) image filtrée par le filtre de Canny	89
Figure 3.16 : (à gauche) Image originale ; (à droite) image filtrée par le filtre de Canny Couleur.....	90
Figure 3.17 : Filtrage à différentes échelles d'une même image (de gauche à droite et de haut en bas) l'image d'origine, l'image avec un niveau d'information élevé (0.5), l'image avec un niveau d'information acceptable (3.5) et l'image avec un niveau d'information faible (5.5). ..	92
Figure 3.18 : Représentation d'une pyramide	92
Figure 3.19 : Schéma général de l'algorithme	94
Figure 3.20 : De gauche à droite : image originale, carte de contours sur la luminance (Value HSV) et carte de contours sur la Teinte (Hue HSV).....	94

Figure 3.21 : Exemples de cartes de contours pour une même image : (de gauche à droite et de haut en bas) ; image originale, images pour un filtrage gaussien différent sigma de 0.5, 3.5 et 5.5.....	95
Figure 3.22 : Cartes de contours pour une même image à chaque niveau de la pyramide	97
Figure 3.23 : Image de fusion des cartes de contours	98
Figure 3.24 : Carte représentant les graines sélectionnées.....	99
Figure 3.25 : Carte de contours finale	99
Figure 3.26 : De gauche à droite et de haut en bas : image originale, et les résultats pour le filtre de Sobel, le filtre de Canny, le filtre Canny couleur et notre pyramide	100
Figure 3.27 : Image originale et les images pour le filtre de Sobel et la méthode pyramidale	101
Figure 3.28 : Image originale et les images pour le filtre de Canny et méthode pyramidale.	101
Figure 3.29 : Image originale et les images pour le filtre de Canny-couleur et notre pyramide	102
Figure 3.30 : Schéma détaillé de la méthode	102
Figure 3.31 : Exemple d'image et sa vérité terrain, les 4 images sont en annexe 3.....	103
Figure 4.1 : Schéma de côté du dispositif expérimental (à gauche) une photographie du dispositif (à droite)	110
Figure 4.2 Répartition de l'âge des malvoyants ayant répondu au questionnaire.....	113
Figure 4.3 Catégorie socio-professionnelles des malvoyants en pourcentage par rapport à chaque groupe.	113
Figure 4.4 Répartition des acuités monoculaires en vision de loin.	114
Figure 4.5 Répartition des malvoyants en fonction de l'atteinte du champ visuel.	115
Figure 4.6 : Taux de réponses pour chaque méthode uniquement pour l'épaisseur de rehaussement le plus proche de l'acuité de chaque malvoyant.....	118
Figure 4.7 : Taux de réponses pour chaque méthode uniquement l'épaisseur de rehaussement le plus proche de l'acuité une fois supérieure de chaque malvoyant	118
Figure 4.8 : Taux de réponses pour chaque méthode uniquement pour l'épaisseur de rehaussement le plus proche de l'acuité une fois inférieure de chaque malvoyant.....	118
Figure 4.9 : Taux de réponses pour chaque méthode uniquement pour l'épaisseur de rehaussement le plus proche de l'acuité de chaque malvoyant.....	119
Figure 4.10 : Taux de réponses pour chaque méthode uniquement l'épaisseur de rehaussement le plus proche de l'acuité une fois supérieure de chaque malvoyant	119
Figure 4.11 : Taux de réponses pour chaque méthode uniquement pour l'épaisseur de rehaussement le plus proche de l'acuité une fois inférieure de chaque malvoyant.....	119
Figure 4.12 : Taux de réponses favorables au rehaussement des contours par rapport aux images sans traitement	120
Figure 4.13 : Comparaison du taux de réponses favorables pour chaque méthode par rapport aux images sources, ou de la 2 ^{ème} méthode du couple par rapport à la 1 ^{ère} méthode du couple de traitements.	122
Figure 4.14 : Taux de préférence de chaque couleur de rehaussement en fonction de l'atteinte du champ visuel.....	123
Figure 4.15 : Nombre de divergences de réponses pour un même couple de couleurs pour les malvoyants du groupe 1 en fonction du type d'atteinte du champ visuel	124
Figure 4.16 : Nombre de divergences de réponses pour un même couple de couleurs pour les malvoyants du groupe 2 en fonction du type d'atteinte du champ visuel	124
Figure 4.17 : Choix de profil de trait en fonction de l'atteinte du champ visuel	125
Figure 4.18 : L'épaisseur du trait (en bleu) en fonction de l'atteinte du champ visuel et l'écart type (en rouge) pour le groupe 1.	126

Figure 4.19 : L'épaisseur du trait (en bleu) en fonction de l'atteinte du champ visuel et l'écart type (en rouge) pour le groupe 2.	126
Figure 5.1 : Exemple de produit de réalité virtuelle : un Trivisio ARVision 3D.....	129
Figure 5.2 : Exemples de produits de réalité augmentée : de gauche à droite ; le masque MG1, le Pro Mobile Display et le Laster See-Thru de Laster Technologies	129
Figure 5.3 : Exemples de produits permettant l'affichage d'information : de gauche à droite ; la Google Glass et le Vuzix M100	130
Figure 5.4 : Sources vidéo possibles pour les dispositifs d'affichage sur tête [158].	132
Figure 5.5 : Quelques exemples de dispositifs optiques de réalité augmentée [35].....	134
Figure 5.6 : Exemples de dispositifs de réalité augmentée de la société Laster Technologies [100, 99, 101, 102]	134
Figure 5.7 : De gauche à droite : l'image originale, l'image incluant une contre distorsion à afficher sur la source vidéo et l'image observée par l'utilisateur à travers le système.	135
Figure 5.8 : Image monoculaire pour un champ de vision augmenté de 40 X 30 degrés et un champ de vision total de 80 X 60 degrés.	136
Figure 5.9 : Image finale binoculaire superposées à une scène : à gauche la situation confortable et à droite une situation inconfortable	137
Figure 5.10 : Images gauche et droite pour une situation confortable (colonne de gauche) et inconfortable (colonne de droite).	138
Figure 5.11 : Schéma du modèle binoculaire	139
Figure 5.12 : Image finale binoculaire superposée à une scène observée.....	141
Figure 5.13 : Images gauche et droite correspondant à une situation confortable du type Figure 5.12	141
Figure 5.14 : Chaîne de calibrage de réalité augmentée	142
Figure 5.15 Illustrations du recalage d'information en réalité augmentée.....	143
Figure 5.16 : Schéma de principe d'un élément photosensible.....	145
Figure 5.17 : Transfert des charges dans une matrice CCD	145
Figure 5.18 : (à gauche) Image sans blooming ; (au centre) avec un blooming vertical ; (à droite) avec un blooming vertical et horizontal.	146
Figure 5.19 : Schéma de principe d'une matrice CMOS	146
Figure 5.20 : (à gauche) Image acquise par une caméra CMOS « rolling shutter » ; (à droite) une caméra « global shutter ».....	148
Figure 5.21 : Interface pour l'ajustement des zones d'intérêts caméra et écran	151
Figure 5.22 : Schéma général de l'application	152
Figure 5.23 : Paramètres utilisateurs	153
Figure 5.24 : Réglage de la position relative de la loupe dans la zone d'affichage	153
Figure 5.25 : Loupe centrée en vert : (à gauche) grossissement de 1 ; (à droite) grossissement de 2.	154
Figure 5.26 : Dispositif embarqué : (à gauche) la carte ; (à droite) le dispositif complet de test (carte + écran + périphériques).	155
Figure 5.27 : Comparaison des temps de traitement en fonction de la plateforme de calcul pour « l'implémentation simplifiée ».	156
Figure 5.28 : Comparaison des temps de traitement en fonction de la plateforme de calcul pour « l'implémentation complète ».	156
Figure A4.1 : Courbes de l'indicateur Précision pour chaque image en fonction du détecteur de contours.	175
Figure A4.2 : Courbes de l'indicateur Rappel pour chaque image en fonction du détecteur de contours.	176
Figure A4.3 : Courbes de l'indicateur F-Mesure pour chaque image en fonction du détecteur de contours.	177

Figure A4.4 : Courbes de l'indicateur Pnd pour chaque image en fonction du détecteur de contours.	177
Figure A4.5 : Courbes de l'indicateur Pco pour chaque image en fonction du détecteur de contours.	178
Figure A4.6 : Courbes de l'indicateur Pfa pour chaque image en fonction du détecteur de contours.	179
Figure A4.7 : Courbes de l'indicateur IMP pour chaque image en fonction du détecteur de contours.	179
Figure A4.8 : Courbes de l'indicateur D4L2 pour chaque image en fonction du détecteur de contours.	180
Figure A4.9 : Courbes de l'indicateur L1 pour chaque image en fonction du détecteur de contours.	181
Figure A4.10 : Courbes de l'indicateur L2 pour chaque image en fonction du détecteur de contours.	181
Figure A4.11 : Courbes de l'indicateur L3 pour chaque image en fonction du détecteur de contours.	181
Figure A4.12 : Courbes de l'indicateur L4 pour chaque image en fonction du détecteur de contours.	182

Remerciements

Je tiens tout d'abord à remercier l'ensemble des personnes avec qui j'ai travaillé, échangé au laboratoire Hubert Curien et qui ont permis de rendre cette thèse enrichissante : les doctorants et membres de l'équipe.

Je remercie tous mes collègues de Laster Technologies et plus particulièrement Stéphane, Gaëlle, Guillaume, Hugues, Lien, Maxime, Merzouk, Rachid, Thierry, Fabrice, Pierre, Thibault, Van-than, Valérian, et Michel.

Je tiens à remercier Jean-Jacques Frayssinet et l'association Rétina France et Vision et Recherche pour leur collaboration lors des essais.

Je tiens à remercier le Docteur Gérard Dupeyron pour m'avoir permis de collaborer avec l'Institut ARAMAV. Je remercie également Michel Dumas pour les échanges et l'organisation d'essais au sein de l'Institut ARAMAV, ainsi que tous les membres de cet Institut.

Je remercie Zile Liu, Éric Dinet, Hubert Konik et Alain Trémeau pour m'avoir permis de réaliser cette thèse et pour m'avoir accompagné tout au long de celle-ci.

Enfin, je remercie également ma famille pour m'avoir soutenu et encouragé dans ce projet.

Introduction

D'après l'organisation mondiale de la santé [126, 2, 125, 134], il y a actuellement dans le monde 285 millions de personnes avec des troubles plus ou moins marqués de leur vision, dont 246 millions de malvoyants. En France, lors d'une étude de l'INSEE réalisée en 1999 [71, 165], il avait été recensé 1 700 000 déficients visuels dont 61 000 aveugles complets. Les malvoyants profonds étaient au nombre de 207 000 personnes. Ils présentaient une vision très faible se limitant à la détection de silhouettes et de formes. Les malvoyants moyens étaient au nombre de 932 000 personnes. Ils présentaient des difficultés pour la reconnaissance des formes, des visages, pour lire et écrire. Les malvoyants légers étaient au nombre de 550 000.

Une partie de ces données a été actualisée en 2011 par l'Institut de la vision. Le Professeur José Alain Sahel et cet institut avançaient 1 500 000 personnes concernées par la DMLA (dégénérescence Maculaire liée à l'âge), 800 000 personnes concernées par le glaucome, entre 400 000 et 500 000 personnes concernées par la rétinopathie diabétique et 40 000 personnes concernées par la rétinite pigmentaire. Ces nouveaux chiffres donnent une estimation entre 2 740 000 à 2 840 000 malvoyants, sans prendre en compte les autres maladies de la vue.

La dégénérescence maculaire liée à l'âge est une atteinte rétinienne évolutive qui affecte la rétine centrale et qui engendre la présence d'un scotome dans le champ de vision de l'observateur. L'observateur ne peut ainsi percevoir une scène qu'avec sa rétine périphérique. La lecture [182, 18, 96], la reconnaissance des visages [26], et plus généralement l'analyse visuelle de toute image sont dès lors perturbées. La DMLA est aujourd'hui la première cause de handicap visuel chez les personnes âgées de plus de 50 ans dans les pays développés [133, 12].

Compte tenu des progrès récents, la création d'aide basée sur des lunettes à réalité augmentée est susceptible de résoudre une partie des problèmes de perception de l'environnement. L'idée sous-jacente est de restituer dans le champ de vision fonctionnelle la partie non perçue de la scène et essentielle à sa compréhension. Ce n'est alors plus le patient qui s'adapte aux aides visuelles, mais le système qui s'adapte aux patients.

Ce concept d'aide se doit d'intégrer un traitement d'images en tenant compte des capacités et des besoins cognitifs des malvoyants, afin de créer une symbiose entre l'image et les mouvements oculaires. Il est à noter que l'être humain n'exploite pas toutes les informations visuelles d'une scène pour interagir avec son environnement.

Aujourd'hui seulement des cannes blanches, des télé-agrandisseurs, et quelques accessoires électroniques permettent aux personnes malvoyantes de se déplacer et de pouvoir travailler en vision de près pour lire un document, ou pour écrire. De plus, des études ont montré qu'il est possible de mettre en forme le texte et certaines données image afin que la personne puisse les percevoir plus facilement et de pouvoir ainsi contourner des problèmes liés à la maladie comme le scotome central ou l'atteinte périphérique du champ visuel. Dans ce cadre, la vision artificielle est devenue un moyen à fort potentiel pour acquérir, modifier ou enrichir l'information visuelle. Ces systèmes semblent pertinents pour la lecture de documents « simples » sans prise en compte ni de la pathologie de la personne, ni de ses problèmes de champ visuel. L'idée générale est simplement d'agrandir l'image, et de modifier ses couleurs d'affichage. Pourtant, les malvoyants auraient besoin que l'image soit enrichie afin d'accentuer les zones informatives en fonction de leur pathologie. Les verrous technologiques à lever sont alors : l'identification des données pertinentes, leur mise en forme et leur restitution.

L'objectif de mes travaux a été de concevoir un outil dédié à l'extraction et à la mise en forme du contenu image pour l'assistance des personnes malvoyantes. Le but final de ce travail est la mise au point d'une aide visuelle basée sur des lunettes informatives adaptées aux malvoyants. Des collaborations ont été établies d'une part, avec l'ARAMAV située à Nîmes (centre de rééducation pour malvoyants, Docteur Dupeyron), non seulement pour définir les besoins opérationnels du système mais également pour les valider, d'autre part, avec l'association Rétina France et son entité Vision & Recherche.

La première partie de ma thèse (chapitres 1 et 2) a été dédiée à l'étude du besoin des personnes déficientes visuelles afin de déterminer les informations, les zones et les contenus image les plus pertinents pour lire et travailler en vision de près, en vision intermédiaire et en vision de loin. Lors de cette phase, une observation du comportement de personnes malvoyantes a été opérée pour déterminer comment les données environnantes sont utilisées pour réaliser ces tâches. Afin de mener à bien cette analyse, j'ai réalisé des entretiens avec des personnes malvoyantes. J'ai également réalisé une immersion au sein de l'ARAMAV afin de mieux appréhender les techniques apprises aux malvoyants pour réaliser les tâches de la vie de tous les jours.

La deuxième partie (chapitres 3 et 4) a été dévolue au développement d'un modèle algorithmique d'extraction de l'information image considérée comme la plus importante, à savoir les contours. J'ai effectué une comparaison des méthodes existantes, l'analyse des performances de celles-ci, leurs avantages et inconvénients. J'ai alors réalisé le

développement d'un outil adapté aux personnes malvoyantes (chapitre 3). Cet outil a vocation à être flexible pour s'adapter, ou être adapté à la situation, et/ou en fonction des besoins du patient. Durant cette phase, j'ai également développé un outil pour restituer, superposer ces informations sur des images fixes. Durant cette deuxième partie, j'ai piloté des essais sur images fixes sur des écrans d'ordinateur (chapitre 4). Cette expérimentation avait pour objectif de mettre en évidence l'apport des contours et les méthodes à utiliser.

La troisième partie (chapitre 5) a consisté au développement d'une solution mobile portable pour les personnes malvoyantes. Pour cela, j'ai développé un modèle permettant d'afficher des informations en binoculaire. L'objectif est d'afficher les informations en monoculaire ou binoculaire en fonction des besoins des patients. J'ai développé un dispositif mobile autonome relié à une paire de lunettes de réalité augmentée à verre transparent et intégrant une caméra.

Il a également été nécessaire de développer un logiciel optimisé pour le calcul temps réel sur ce système. Ce développement a également nécessité l'intégration d'une interface permettant aux patients d'interagir avec le dispositif.

Dans ce manuscrit, je vais commencer par présenter des notions liées au système visuel humain afin de mieux comprendre les besoins des malvoyants. Je vais ensuite présenter les analyses quantitatives et qualitatives réalisées sur les besoins des malvoyants. Je vais ensuite présenter les développements de traitements d'images que j'ai réalisés, ainsi que les évaluations qui ont été effectuées par les malvoyants. Pour finir, je vais présenter les différents éléments d'un dispositif, de réalité augmentée, basé sur mes travaux de recherche, qui a été développé par Laster Technologies, pour répondre aux besoins spécifiques des personnes malvoyantes.

Synthèse

Afin de répondre aux besoins des malvoyants, ma thèse a été de concevoir un outil dédié à l'extraction et à la mise en forme du contenu image pour l'assistance des personnes malvoyantes. Le but final de ce travail est la mise au point d'une aide visuelle basée sur des lunettes informatives adaptées aux malvoyants.

Il est nécessaire :

- De définir les signes visuels importants pour les malvoyants à rehausser :
 - Quels sont les indices visuels les plus utiles aux malvoyants ?
 - Existe-t-il un lien entre leurs pathologies et cette réponse ?
 - Quels sont leurs usages des aides visuelles ?
 - Ces aides visuelles répondent-elles à leurs besoins ?
- D'évaluer les méthodes existantes pour détecter ces indices visuels :
 - Quel est l'apport du rehaussement des contours pour les malvoyants ?
 - Quel est l'apport de ce rehaussement avec les méthodes existantes ?
- De développer et valider une nouvelle méthode de détecter si nécessaire :
 - Comment réaliser un détecteur de contours adapté aux besoins des malvoyants ?
 - Quel est l'apport du rehaussement dans ce cas ?
- De définir les moyens de rehausser ces indices visuels pour apporter une aide aux malvoyants :
 - Comment (taille, couleurs, forme) doit-on rehausser les contours ?
 - Quel lien existe-t-il entre le rehaussement à faire et les différentes atteintes de la vision (champ visuel, acuité, ...) ?
- D'identifier les composants et matériels nécessaires pour la réalisation de l'aide visuelle finale :
 - Quel type de dispositif d'affichage utiliser ?
 - Quel type de caméra utiliser ?
 - Quel type de système de calcul permet de répondre à ce besoin ?

1. Le système visuel et la basse vision

Ce chapitre présente le système visuel humain ainsi que les principales pathologies entraînant une perte plus ou moins marquée de la vision du patient. Je vais dans un premier temps introduire le principe de formation d'image sur la rétine. Je vais ensuite présenter le fonctionnement de la rétine et surtout des champs récepteurs. Je vais ensuite lister et expliquer un certain nombre de pathologies pouvant occasionner une déficience visuelle : baisse de l'acuité et/ou modification du champ visuel, baisse de la perception des contrastes et/ou des couleurs. Enfin, je vais présenter le cortex visuel et le fonctionnement de la vision binoculaire. Ce chapitre intègre les définitions nécessaires pour la bonne compréhension de la suite du document.

1.1. L'œil

1.1.1. La formation de l'image

L'œil est un l'organe qui permet au système visuel humain de collecter les signaux lumineux et de fournir des influx nerveux en lien avec ces signaux. Ces signaux nerveux sont ensuite transférés par les voies optiques au cortex visuel. La perception visuelle est la représentation mentale des signaux lumineux acquis par les yeux, par le cortex visuel.

L'œil [95] est composé d'éléments optiques tels que la cornée, l'humeur aqueuse, l'iris, le cristallin, le corps vitré, et la rétine, cf. Figure 1.1.

L'humeur aqueuse est située entre la cornée et le cristallin. Le corps vitré est situé entre le cristallin et la rétine.

La cornée est la face avant de l'œil. Elle constitue le premier élément optique permettant de focaliser les signaux lumineux sur la rétine. Sa puissance optique (vergence) ne varie pas.

L'iris est situé entre la cornée et le cristallin. Il est quasiment collé à la face avant du cristallin. Son rôle est de limiter la quantité de lumière et de réduire les aberrations optiques pour la formation d'une image rétinienne nette. Il joue le rôle de diaphragme optique.

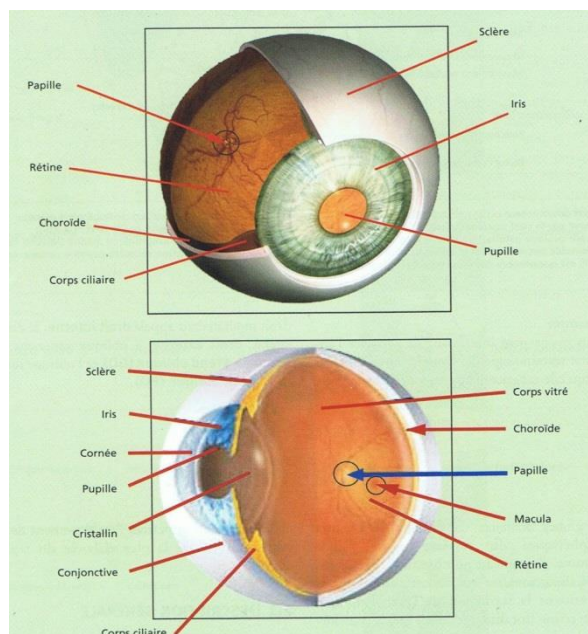


Figure 1.1 : Représentation de l'œil¹

Le cristallin est le deuxième élément optique de l'œil. Il a une puissance optique (vergence) variable. Son rôle est de focaliser l'image des signaux optiques issue de la cornée sur la rétine quelles que soient leurs distances de provenance. L'adaptation de la vergence du cristallin est réalisée par la contraction des muscles fixant celui-ci dans l'œil. Cela a pour effet de modifier les courbures de sa face avant et de sa face arrière. Ce phénomène est appelé accommodation ; il est mis en œuvre de manière réflexe par le système visuel. Afin de voir un objet à l'infini un œil emmétrope (ne présentant aucun défaut optique) ne sollicite pas son accommodation. Lorsque l'objet n'est plus à l'infini, l'image de celui-ci, n'est plus localisée sur la rétine, mais derrière celle-ci. Afin de refocaliser l'image de cet objet sur la rétine, il faut augmenter la vergence de l'œil. Pour voir un objet rapproché net, il faut solliciter son accommodation d'une valeur inversement proportionnelle à la distance de l'objet. Plus l'objet est proche plus l'accommodation nécessaire sera forte. Avec l'âge, le cristallin peut s'opacifier dans ce cas, on parle de cataracte. De même avec l'âge, le cristallin perd de sa plasticité, il devient un peu plus dur et les muscles le contrôlant sont moins actifs. La capacité d'accommodation baisse donc avec l'âge, lorsque l'accommodation devient trop faible pour regarder à une distance rapprochée des objets de manière confortable, on parle de presbytie.

¹ Eric Beaubert, Franck Pariguet, et Stéphane Taboulot. *Manuel de l'opticien*. ISBN 2-224-02813-X. Editions Maloine, 2005

1.1.2. Critères d'analyse de la vision

Avant de présenter un certain nombre de troubles de la vision qu'il s'agisse de troubles optiques, de troubles des voies optiques, ou des anomalies rétinienne, je vais présenter quelques critères, utilisés couramment par les opticiens, les ophtalmologistes et les orthoptistes, permettant de qualifier la vision d'une personne. La liste de critères présentés ici, n'est pas exhaustive. On peut citer :

- l'acuité visuelle
- la sensibilité au contraste
- le champ visuel

1.1.2.1. L'acuité visuelle

L'acuité visuelle est la capacité à voir distinctement les détails d'un objet dans le but de reconnaître celui-ci. Cette définition de l'acuité correspond à celle de l'acuité morphoscopique, c'est-à-dire l'acuité mesurée lorsque l'on reconnaît des lettres différentes ou des formes, sur une ligne. Il existe également l'acuité du minimum visible : la capacité à détecter un élément ponctuel et l'acuité du minimum séparable : la capacité à séparer deux points sur une même ligne.

L'acuité morphoscopique se caractérise en fonction de l'Angle Minimal de Résolution : AMR. Il s'agit de la taille angulaire du plus petit détail perceptible.

$$\text{acuité visuelle} = \frac{1}{\text{AMR (en minute d'angle)}}$$

Les optotypes sont des lettres, chiffres ou symboles normalisés. Ils sont de hauteur et largeur identiques, généralement égales à 5 fois l'AMR. Ces symboles permettent de réaliser différentes échelles d'acuité. Il existe notamment

- l'échelle décimale (1/10, 2/10 ... 10/10) : elle est très répandue. Elle permet de mesurer de bonnes acuités ($>5/10^{\text{ème}}$) mais ne présente que peu de pas pour les faibles acuités ($<5/10^{\text{ème}}$).
- l'échelle aux inverses (1/10, 1/9, ... 1/1) : cette échelle est très adaptée pour la mesure des faibles acuités, elle présente un grand nombre de pas dans les faibles acuités.

- l'échelle rationnelle : il s'agit d'une combinaison des échelles aux inverses et décimales afin de bénéficier à la fois de la précision pour les bonnes acuités et les faibles acuités.
- l'échelle logarithmique (log MAR) : cette échelle présente un pas constant entre chaque ligne, cf. Figure 1.2. Cette échelle est donc parfaitement adaptée à la mesure d'une bonne ou d'une faible acuité. Il est possible de convertir une acuité décimale en acuité log MAR et vice versa.

$$acuité\ visuelle = 10^{-(mesure\ logMAR)}$$

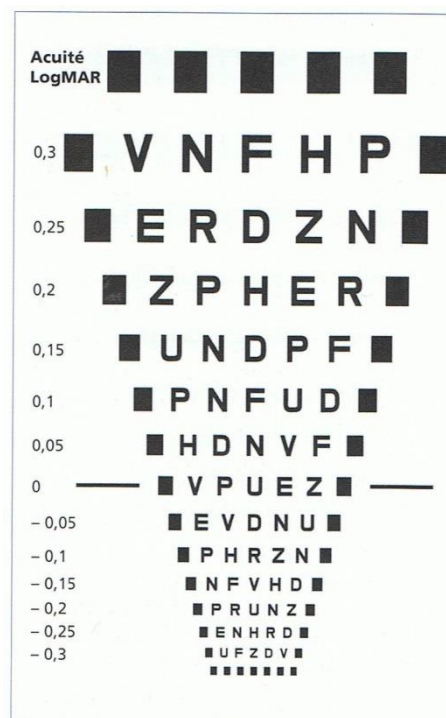


Figure 1.2 : Échelle logarithmique log MAR²

1.1.2.2. Le champ visuel

Le champ visuel définit la zone de l'espace réel perceptible par un œil fixe. Pour un œil sain, le champ visuel normal, par rapport à l'axe optique de l'œil, est de 60° vers le nez, 100° vers la tempe, 65° vers le haut et 75° vers le bas, cf. Figure 1.3. En binoculaire, le champ visuel peut atteindre 200° à l'horizontale. Dans ce cas, la partie commune aux deux yeux est de l'ordre de 60°.

² Éric Beaubert, Franck Pariguet, et Stéphane Taboulot. *Manuel de l'opticien*. ISBN 2-224-02813-X. Editions Maloine, 2005

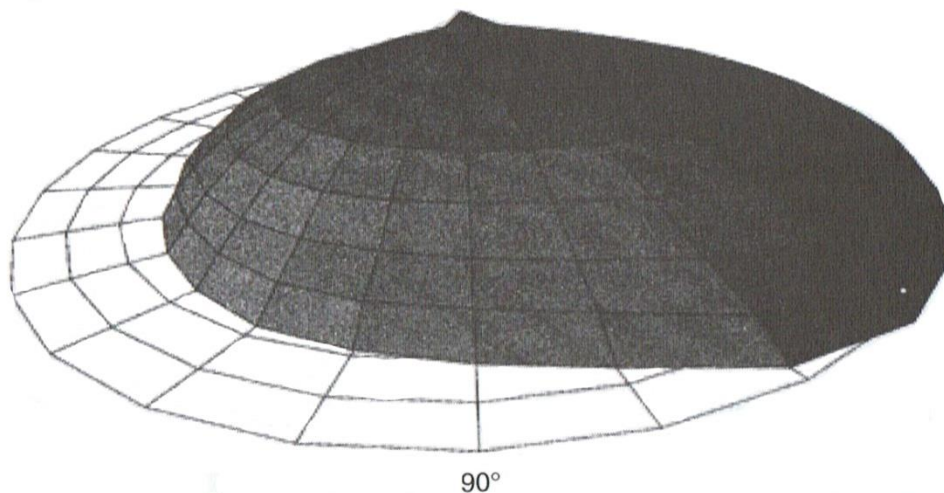


Figure 1.3 : Acuité en fonction du champ visuel³

Comme cela est présenté, dans la section sur la rétine, la zone centrale du champ visuel présente une très bonne acuité, car elle correspond à la fovéa. Le reste du champ visuel présente une acuité beaucoup plus faible et permet surtout la perception du mouvement.

Le champ visuel peut être modifié par la présence de zones non-fonctionnelles ou moins fonctionnelles en fonction de la pathologie pouvant affecter l'œil.

1.1.2.3. La sensibilité au contraste

Le contraste quantifie la différence maximale entre la valeur la plus sombre et la plus lumineuse entre deux éléments. Il est généralement défini par la formule :

$$\text{Contraste} = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}$$

Avec I_{\max} , la valeur maximale de l'élément le plus clair à comparer et I_{\min} la valeur minimale de l'élément le plus sombre.

Une mesure de la plage de contraste peut permettre de qualifier certaines aptitudes visuelles de la personne, pour détecter certaines pathologies telles que le glaucome de manière précoce.

Le système visuel n'a pas la même sensibilité en fonction de la fréquence spatiale et en fonction du contraste, cf. Figure 1.4. Comme le montre les courbes de la Figure 1.4, la sensibilité de l'œil en fonction du contraste évolue avec l'âge.

³ Caroline Kovarski. *L'opticien-lunetier guide théorique et pratique*. ISBN 2-7430-0651-X. Edition Lavoisier, 2005 et extrait de : Atlas d'histologie (1re édition française) Sobotta J., Welsch U. © Lavoisier, 2002 (pour l'édition française) © Urban & Fischer, 2002 (pour l'édition originale allemande)

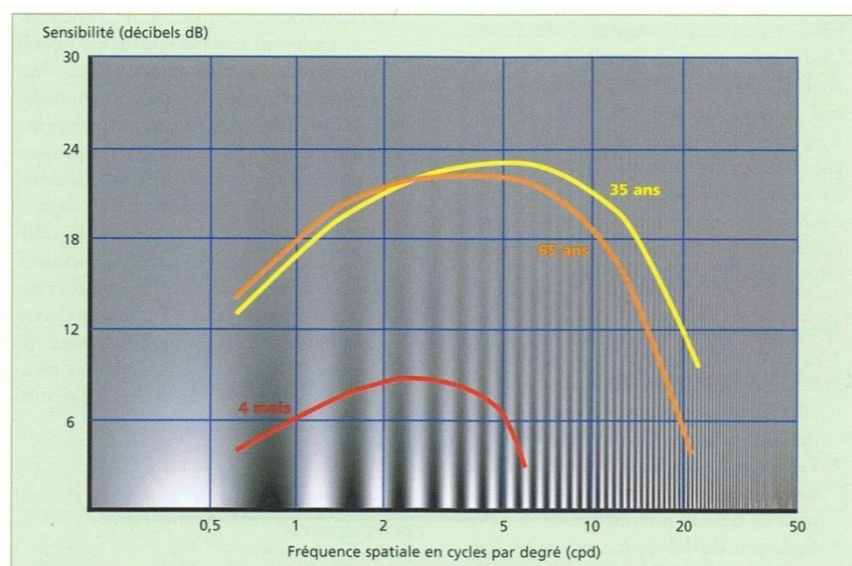


Figure 1.4 : Courbe de sensibilité aux contrastes⁴

Cette mesure peut être réalisée de plusieurs manières, soit en utilisant des échelles d'acuité avec plusieurs valeurs de contrastes différentes soit en faisant varier la valeur du contraste entre deux éléments lumineux et en analysant quand le patient peut détecter ces variations.

1.1.3. Amétropies

Les amétropies sont des anomalies de la vergence de l'œil ou de la forme de globe oculaire. Pour un individu, elles peuvent être différentes pour l'œil gauche et l'œil droit. Ces amétropies peuvent être une anomalie de la courbure d'une des faces de la cornée, d'une des faces du cristallin, une variation de la forme ou de la longueur de l'œil. En d'autres termes, une amétropie est la source de la formation d'une image mal focalisée sur la rétine ou déformée. On peut citer, l'hypermétropie, la myopie, l'astigmatisme. Les amétropies peuvent également être dues au vieillissement du cristallin, on parle alors de presbytie. Ces amétropies sont très souvent compensables par des systèmes optiques tels que les verres correcteurs. Lorsque les amétropies ne sont pas compensées, elles peuvent impacter, gêner le confort visuel lors de l'utilisation des dispositifs de réalité augmentée.

La forte myopie est la seule amétropie qui peut être considérée comme une basse vision, car elle peut engendrer le même type de baisse de l'acuité et être source de troubles rétinien.

⁴ Eric Beaubert, Franck Pariguet, et Stéphane Taboulot. *Manuel de l'opticien*. ISBN 2-224-02813-X. Editions Maloine, 2005

1.1.3.1. La myopie

La myopie est une amétropie due, soit à une vergence trop élevée de l'œil, soit un œil anormalement long. Dans les deux cas, l'image est focalisée avant la rétine, l'image rétinienne est donc floue. Pour refocaliser cette image, il faudrait « désaccommoder » ce qui est impossible. Une personne myope peut donc voir des objets nets en vision de près, cf. Figure 1.5. Le verre correcteur permet au patient lorsque son accommodation est nulle de voir net à l'infini.

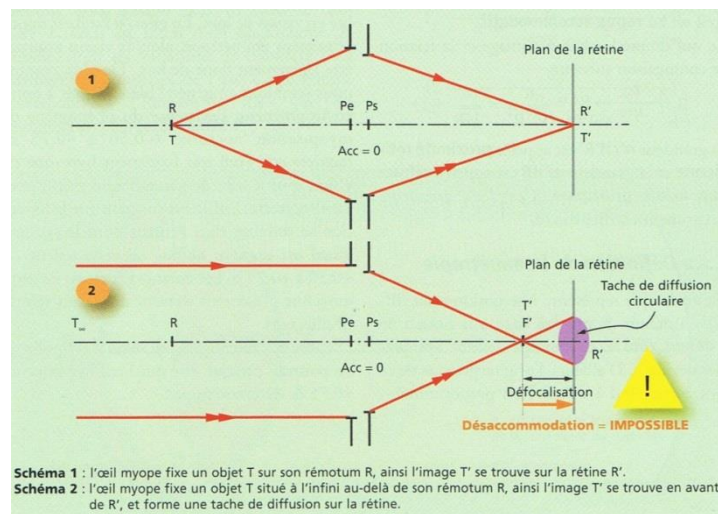


Figure 1.5 : Schéma d'un œil myope non corrigé⁵

La forte myopie, généralement supérieure à 6 dioptries^{6,7,8} peut, dans certains cas, être considérée comme de la basse vision. De par la forme trop allongée de l'œil, la rétine est fragilisée et est donc plus exposée à la néo-vascularisation et au trou maculaire. Cela peut entraîner une perte de la vision centrale et/ou une baisse de l'acuité. Cette amétropie peut être prise en compte dans les pathologies de la basse vision [45].

1.1.3.2. L'hypermétropie

Cette amétropie est l'opposé de la myopie, soit l'œil est trop petit, soit la vergence de l'œil est trop faible. Dans les deux cas, l'image est floue sur la rétine et l'image nette se forme derrière de la rétine, cf. Figure 1.6.

⁵ Eric Beaubert, Franck Pariguet, et Stéphane Taboulot. *Manuel de l'opticien*. ISBN 2-224-02813-X. Editions Maloine, 2005

⁶ <http://www.gatinel.com/recherche-formation/myopie-definition-mecanismes-epidemiologie-facteurs-de-risques/classification-de-la-myopie/>

⁷ <http://www.braille.be/fr/documentation/pathologies-visuelles/myopie-forte>

⁸ <http://www.ophtalmologie.fr/myopie.html>

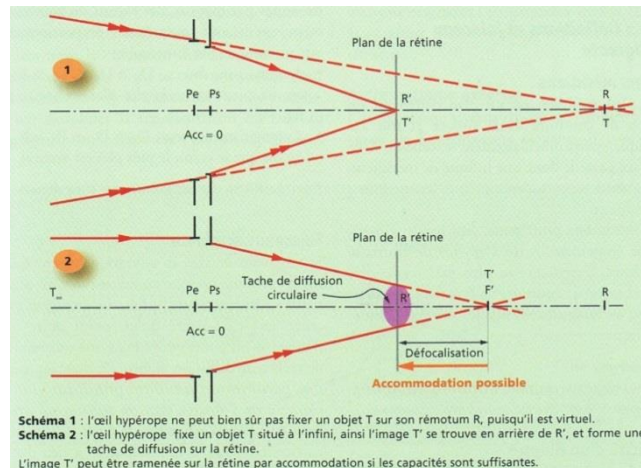


Figure 1.6 : Schéma d'un œil hypermétrope non corrigé⁹

Afin de voir nette une image à l'infini, le patient peut focaliser l'image sur la rétine en sollicitant son accommodation. Le patient pourra faire de même en vision de près. Un hypermétrope de 2 dioptries devra solliciter 5 dioptries pour regarder un objet à 33 cm au lieu de 3 dioptries comme pour un emmétrope. Cette amétropie peut être compensée par un verre correcteur qui augmente la vergence du système verre œil.

Lorsque la capacité d'accommoder diminue avec l'âge, il devient plus difficile de compenser l'hypermétropie en accommodant. Dans ce cas, cet amétrope verra net les objets à l'infini et verra flou les objets proches.

1.1.3.3. L'astigmatisme

L'astigmatisme est une amétropie pouvant concerner la myopie ou l'hypermétropie. Dans le cas de cette amétropie, l'œil ne possède pas une vergence uniforme, mais présente deux vergences suivant deux méridiens différents. Cela peut être dû à la forme de l'œil ou à la forme d'une ou des faces du cristallin ou de la cornée. Cette amétropie peut être compensée par des verres toriques afin de compenser les vergences suivant chaque méridien. Les tailles des images rétinienne avec corrections optiques peuvent être légèrement différentes entre les deux yeux, appelées aniséïconie ; cela peut poser des troubles de la vision binoculaire. Une grande aniséïconie entre les deux images peut entraîner des troubles pour l'appariement des points entre les images rétinienne gauche et droite.

⁹ Eric Beaubert, Franck Pariguet, et Stéphane Taboulot. *Manuel de l'opticien*. ISBN 2-224-02813-X. Editions Maloine, 2005

1.1.3.4. La presbytie

Nous avons mentionné le réflexe d'accommodation. Les capacités d'accommodation diminuent avec le temps à cause du vieillissement du cristallin. Le patient ne peut plus ou très difficilement accommoder pour voir net en vision de près. Il est nécessaire de compenser cette amétropie comme l'hypermétropie en compensant le manque d'accommodation grâce à un verre correcteur. Il est alors souvent proposé soit des verres multifocaux soit des verres progressifs, afin de combiner d'une part la correction pour la vision de loin et d'autre part, les corrections contenant le complément nécessaire pour voir net un objet en vision de près.

1.1.4. La rétine

1.1.4.1. Rôle et constitution

La rétine est la partie photosensible de l'œil, c'est cette partie qui convertit les signaux lumineux en influx nerveux et les transmet par les voies optiques au cortex visuel.

La rétine est constituée entre autres de l'épithélium pigmentaire, de la couche de photorécepteurs, des cellules horizontales, des cellules bipolaires, des ganglionnaires et des fibres du nerf optique, cf. Figure 1.7.

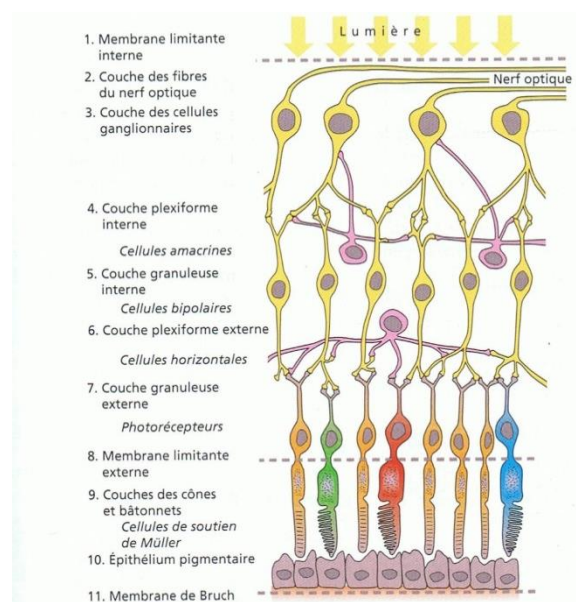


Figure 1.7 : Organisation schématique de la rétine¹⁰

¹⁰ Eric Beaubert, Franck Pariguet, et Stéphane Taboulot. *Manuel de l'opticien*. ISBN 2-224-02813-X. Editions Maloine, 2005

La rétine comprend un grand nombre de photorécepteurs : 6 à 7 millions de cônes et 130 millions de bâtonnets.

Les cônes et les bâtonnets ne sont pas distribués de manière uniforme dans la rétine [46]. La fovéa, zone très petite d'environ, 500 μm de diamètre, au centre du regard, est la zone la plus dense en cônes et est dépourvue de bâtonnet. La macula est une zone plus large, une ellipse de 3 mm par 1,5 mm, contenant la fovéa, elle ne contient presque que des cônes et très peu de bâtonnets. Cette zone de la rétine permet la vision des détails et des couleurs. La zone de vision périphérique contient-elle beaucoup plus de bâtonnets et beaucoup moins de cônes, cf. Figure 1.8. Cette zone de la rétine, la plus grande, permet la perception des mouvements.

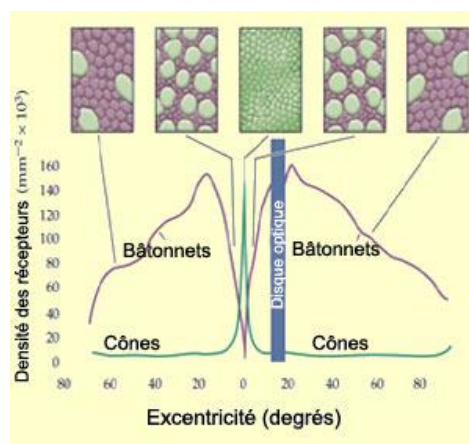


Figure 1.8 : Distribution spatiale des cônes et des bâtonnets au sein de la rétine¹¹

Les cônes ne possèdent pas tous la même sensibilité fréquentielle. La rétine comprend 3 types de cônes, chacun est sensible à une bande spectrale différente, cf. Figure 1.9.

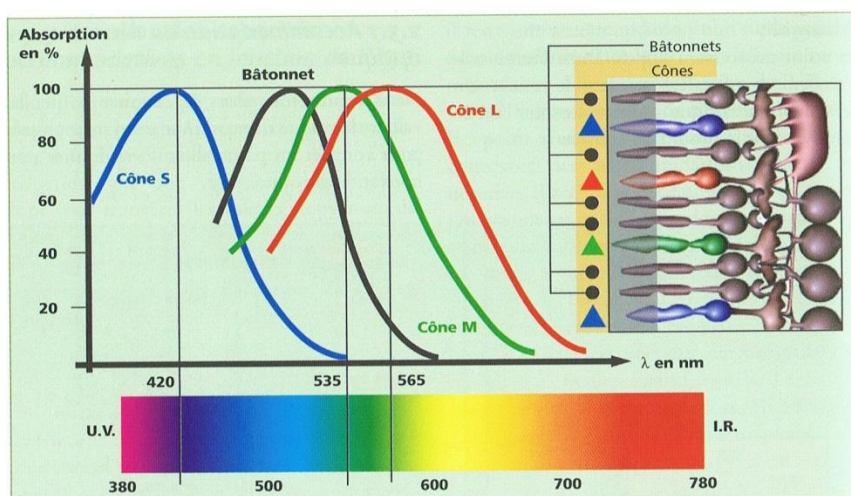


Figure 1.9 : Sensibilité spectrale des cônes et bâtonnets¹²

¹¹ Mc Gill university <http://lecerveau.mcgill.ca/>

1.1.4.2. Champ récepteur

Les influx nerveux provenant des photorécepteurs sont regroupés et compressés par les différentes couches de la rétine, cf. Figure 1.7, les cellules bipolaires, horizontales et ganglionnaires dans le but de simplifier et encoder l'information à transmettre [1]. Nous n'allons pas présenter le rôle et le fonctionnement de chacune de ces cellules, mais nous allons présenter la notion de champ récepteur.

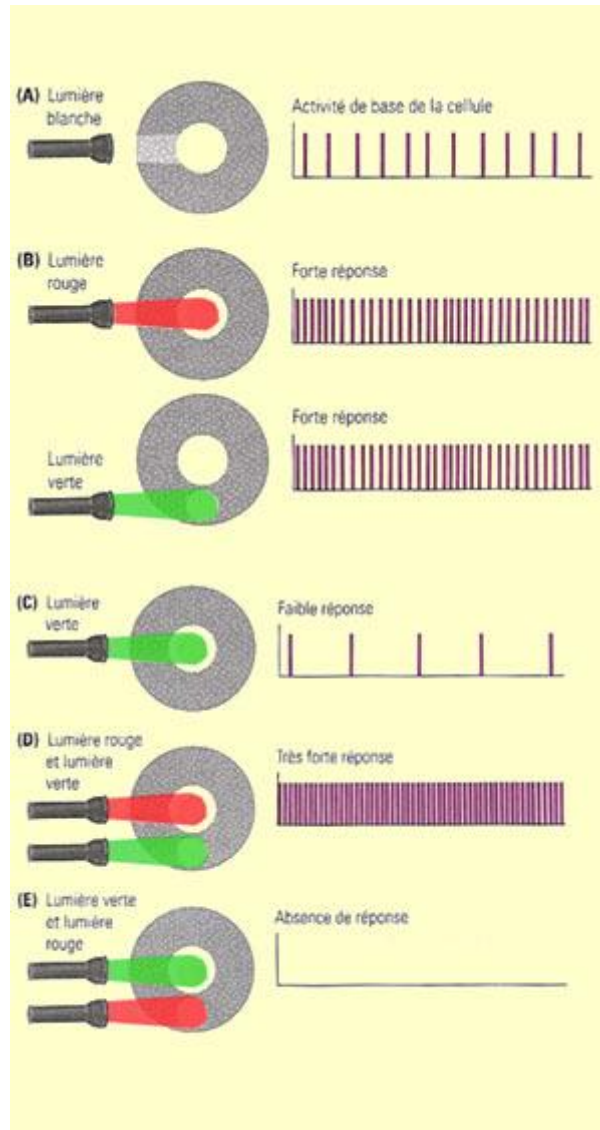


Figure 1.10 : Fonctionnement d'un champ récepteur¹³

¹² Eric Beaubert, Franck Pariguet, and Stéphane Taboulot. *Manuel de l'opticien*. ISBN 2-224-02813-X. Editions Maloine, 2005

¹³ Mc Gill university <http://lecerveau.mcgill.ca/>

Le champ récepteur est une zone d'analyse de la rétine. Chaque champ récepteur fournit une réponse prenant en compte tous les photorécepteurs qui lui sont connectés. La taille de chaque champ récepteur varie en fonction de sa localisation sur la rétine. Les champs récepteurs ont une réponse qui varie en fonction des influx nerveux en leur centre et en leur périphérie, cf. Figure 1.10. Les champs récepteurs regroupent les influx nerveux de plusieurs photorécepteurs de la rétine, cela permet de transmettre un seul influx nerveux qui code en fonction de sa fréquence plusieurs cas possibles dans un seul champ récepteur. Les champs récepteurs sont fortement sensibles aux variations temporelles et spatiales des informations.

Afin de réduire la quantité d'informations à transmettre au cortex visuel, les champs récepteurs sont plus sensibles aux variations dans l'image. Ces champs permettent donc de compresser les informations à transmettre en conservant uniquement les variations, les changements brusques dans l'image. Les contours sont la délimitation entre des éléments de nature suffisamment différente. De par leur fonctionnement, les champs récepteurs sont très stimulés en présence de contours. De par la présence des champs récepteurs, la rétine transmet beaucoup d'informations relatives aux contours dans l'image observée.

1.1.5. Atteintes de la rétine

Nous allons dans cette partie, présenter les principales maladies de la vue provoquant la basse vision, soit en générant un trouble léger de la vision soit un trouble plus fort pouvant entraîner la cécité complète de la personne. Les chiffres indiqués correspondent aux chiffres pour la France, provenant de l'étude HID réalisée en 1999 par l'INSEE [71, 165]. Dans le chapitre 2, je vais présenter ces chiffres plus en détail ainsi que les estimations actuelles de ces chiffres réalisées par différents organismes. Nous ne pouvons réaliser une liste exhaustive des atteintes de la rétine et du système visuel, car il existe un grand nombre d'atteintes possibles et un très grand nombre de variantes possibles pour chacune de ces atteintes. Je vais seulement présenter dans cette partie les principales atteintes rencontrées lors de mes travaux de recherche. Ces atteintes de la rétine ne sont pas réversibles, les traitements actuels permettent de stopper ou de ralentir la progression. Les cellules de la rétine une fois abimées ne peuvent pas être remplacées ou renouvelées de manière naturelle.

1.1.5.1. La dégénérescence maculaire liée à l'âge

La Dégénérescence maculaire liée à l'âge ou DMLA [16] touchait 1 % de la population soit 650 000 personnes en 1999. Cette pathologie pourrait concerner 30 % des personnes âgées d'au moins 75 ans. Cette maladie peut exister sous deux formes :

- **Forme atrophique:** cette forme de la DMLA concerne 90 % des personnes atteintes. Ces patients conservent une acuité visuelle entre 2/10 et 3/10^{ème}. Elle est due à une atrophie progressive d'une partie des cellules de la rétine, de la choroïde, cf. Figure 1.1, et de l'épithélium pigmentaire, cf. Figure 1.4. Cette atrophie provoque une dégradation progressive de la vision de la personne. Elle se caractérise par la présence de drusens : déchets non évacués de la rétine, dit « durs » à la surface de celle-ci.
- **Forme exsudative:** cette forme de la DMLA concerne 10 % des patients atteints. Cette version est plus dramatique car l'acuité visuelle chute rapidement à 1/100^{ème}. Elle se caractérise par la présence de drusens dit « mous » : drusens plus grands que les drusens dit « durs ». Cette présence entraîne une sur néo-vascularisation de la rétine, cela induit une modification de la surface de la rétine et une chute de la vision [17].

Dans les deux cas cette pathologie affecte le champ visuel avec un scotome, zone peu ou pas fonctionnelle dans le champ visuel, cf. Figure 1.11.



Figure 1.11 : (à gauche) Image d'une vision normale ; (à droite) d'une vision atteinte de DMLA ¹⁴

Dans les deux cas, il y a une perte de la vision centrale avec un maintien de la vision périphérique.

1.1.5.2. Le glaucome

Il existe deux formes de glaucome [16] : le glaucome aigu par fermeture de l'angle et le glaucome chronique à angle ouvert. Ils se caractérisent souvent par un rétrécissement du

¹⁴ National eye institute <http://www.nei.nih.gov/>

champ visuel jusqu'à la cécité, cf. Figure 1.12. Cela est dû à une surpression de l'humeur aqueuse. Cela va occasionner des troubles, pressions réduisant la quantité d'informations transmises, au niveau de la papille optique, zone de la rétine dépourvue de photorécepteur, zone également appelée tâche aveugle, et permettant la connexion du nerf optique. Le glaucome occasionne une baisse de l'acuité s'il n'est pas traité rapidement.



Figure 1.12 : (à gauche) Image d'une vision normale; (à droite) d'une vision atteinte de glaucome
15

1.1.5.3. La rétinopathie diabétique

La rétinopathie diabétique : cette pathologie [16] concerne environ 1 % de la population soit 400 000 à 500 000 personnes. Elle se caractérise par la perte de la vision centrale avec un maintien de la vision périphérique, cf. Figure 1.13. Elle est en lien avec le diabète du patient. Elle provoque des troubles de l'irrigation sanguine de la rétine et peut conduire à la cécité.



Figure 1.13 : (à gauche) Image d'une vision normale ; (à droite) d'une vision atteinte de rétinopathie diabétique ¹⁶

1.1.5.4. La rétinite pigmentaire

La rétinite pigmentaire [16] est souvent héréditaire. Elle se caractérise par une atrophie au début des bâtonnets puis des cônes de la rétine du patient. Cette pathologie concernait environ 35 000 personnes en 1999. Elle se caractérise par une cécité nocturne, un rétrécissement du

¹⁵ National eye institute <http://www.nei.nih.gov/>

¹⁶ National eye institute <http://www.nei.nih.gov/>

champ visuel, pouvant atteindre 2, 3° pour atteindre une vision tubulaire, vision similaire à la celle possible à travers un tube (tunnel, fusil de canon), cf. Figure 1.14. En fonction de l'âge, les patients peuvent conserver une acuité leur permettant de lire. L'acuité et le champ visuel diminuent progressivement avec l'évolution de la pathologie.



Figure 1.14: (à gauche) Image d'une vision normale ; (à droite) d'une vision atteinte de rétinite pigmentaire ¹⁷

1.2. Le système visuel

1.2.1. Introduction

Le système visuel humain regroupe les deux yeux, le cortex visuel ainsi que tous les éléments permettant d'échanger les informations entre les yeux et le cortex visuel.

Les influx nerveux provenant des cônes et des bâtonnets sont compressés, codés par les cellules de la rétine afin de passer d'environ 130 millions de signaux à environ 1 million en sortie de l'œil (chaque nerf optique possède environ 1 million d'axones).

Les voies visuelles permettent d'acheminer ces influx nerveux au cortex visuel, pour cela, les nerfs optiques gauches et droits sont regroupés au niveau du chiasma optique. Les fibres nerveuses sont alors séparées de manière à transmettre au côté gauche du cortex visuel, les fibres transportant les informations du champ visuel droit. De la même manière les fibres nerveuses transportant les influx nerveux issus du champ visuel gauche sont connectées au cortex visuel droit.

Les fibres nerveuses passent par les bandelettes optiques dans les corps genouillés latéraux pour en ressortir par les radiations optiques et ainsi rejoindre les aires primaires du cortex visuel. Le cortex visuel contient un certain nombre d'aires visuelles, celles-ci remplissent une tâche bien définie dans l'analyse des influx nerveux afin de donner à notre cerveau une

¹⁷ National eye institute <http://www.nei.nih.gov/>

perception de la scène. Il existe notamment des aires visuelles dorsales et temporales. Les aires dorsales sont davantage dédiées à la localisation des objets dans l'espace environnant, alors que les aires temporales sont davantage dédiées à la reconnaissance des objets.

Le système visuel humain est un système binoculaire, cela étant, je vais donc tout d'abord présenter quelques informations complémentaires relatives à la vision monoculaire. Dans un deuxième temps que je vais présenter sur le principe de la vision binoculaire.

1.2.2. Vision monoculaire

La vision monoculaire, réalisée à partir de chaque œil, permet au système visuel une certaine perception de la profondeur par l'analyse et l'interprétation d'un certain nombre d'indices visuels. L'occlusion ou le masquage partiel des objets permet de les situer les uns par rapport aux autres. L'analyse de leurs tailles respectives par rapport aux tailles habituelles de ces objets, la perspective sous laquelle les objets sont perçus, et la parallaxe permettent de donner une certaine taille aux objets. Enfin l'analyse des ombres et des effets de lumière permettent d'affiner la perception de la profondeur, de la perspective de la scène en utilisant uniquement des informations monoculaires.

1.2.3. Vision binoculaire

Il est intéressant pour la suite de cette partie d'analyser le mécanisme de la vision binoculaire [16]. Ce mécanisme est important à prendre en considération pour l'usage de dispositifs de réalité augmentée et la compréhension de l'impact de certains troubles de la vision sur la perception de la réalité par le patient.

Un certain nombre des informations de cette partie sera utile dans le chapitre 5 pour la partie concernant les dispositifs de réalité augmentée.

Le système visuel humain comprend deux yeux. En temps normal, nous utilisons nos deux yeux afin de percevoir l'environnement qui nous entoure. Le champ visuel binoculaire peut atteindre 160° voire 200° par 135° [62]. Cette vision comprend 3 niveaux : la vision simultanée d'une même scène par les deux yeux, la fusion de ces images, et la stéréoscopie. L'individu doit bénéficier du premier niveau afin d'espérer atteindre le deuxième niveau, ce deuxième niveau ne peut pas être présent sans le premier niveau. Il en va de même pour le troisième niveau qui nécessite la présence du premier et du deuxième niveau.

Le premier niveau est la vision simultanée ; il est requis pour la vision binoculaire et la stéréoscopie. Il s'agit de la perception d'une image avec chaque œil de manière simultanée. Cela implique l'usage d'une même accommodation et d'une même convergence. Il est nécessaire de prendre en compte la convergence (lien entre l'écart pupillaire (distance entre les centres des deux yeux) et la distance d'observation). La convergence est liée à l'accommodation, il peut être nécessaire de compenser l'amétropie de la personne par une correction ophtalmique.

La figure 1.15 représente en abscisse, la convergence et en ordonnées l'accommodation, en dioptries. La ligne « LNAC » représente la relation naturelle entre l'accommodation et la convergence en vision binoculaire normale. Il est seulement possible d'utiliser un tiers de la zone de vision nette pour une vision binoculaire confortable.

La zone entre les lignes bleues la zone de vision binoculaire nette, en dehors de cette zone, la vision binoculaire n'est pas ou mal existante. La largeur de cette zone varie d'une personne à l'autre. Lorsque l'on s'écarte de plus d'un tiers dans cette zone de vision nette, on entre alors dans la zone des phories (écart important entre l'accommodation et la convergence, il ne faut plus respecter la liaison naturelle), la vision binoculaire lors d'un usage prolongé peut devenir inconfortable (mal de tête, vision double) car l'on sollicite la zone de réserve de la personne (zone dans laquelle la personne peut fusionner l'image mais en réalisant un effort important).

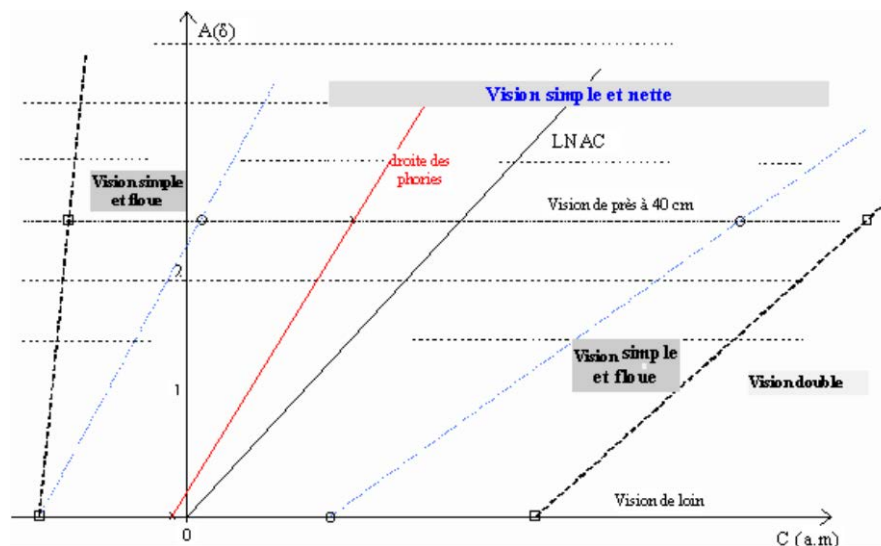


Figure 1.15 : Relation accommodation A (en dioptries)-convergence (C en angle métrique) du système visuel

Le deuxième niveau est la fusion : il s'agit de la fusion des images gauche et droite en une seule et même image. La partie gauche de l'image gauche est placée à gauche, la partie droite

de l'image de droite est placée à droite et au centre est placée une fusion des parties communes aux deux images. Il est possible de définir des zones dans lesquelles la vision binoculaire est possible en fonction de l'accommodation et de la convergence. Les zones du cortex visuel regroupant les mêmes zones du champ visuel binoculaire pour chacun des yeux sont appelées aire de Panum. Chacune de ces aires regroupe la partie de l'image rétinienne gauche et droite correspondant à une même zone ou au même point. En d'autre terme, il s'agit d'aires réalisant l'appariement des champs récepteurs de l'œil gauche avec les champs récepteurs de l'œil droit. L'horoptère est la zone, en respectant la relation accommodation convergence, pour laquelle, il est possible de réaliser la fusion des images. Cette zone peut être mesurée de manière expérimentale, on obtient alors l'horoptère expérimental, cf. Figure 1.16. Il délimite les zones dans lesquelles les objets doivent être observés afin que leurs images soient dans des aires de Panum correspondantes dans le cortex visuel. Lorsque la personne fixe le point O, son accommodation et sa convergence respecte la relation naturelle accommodation-convergence afin que ces lignes de regards passent par ce point. Dans ce cas, les images Md et Mg du point M se forment sur une aire de Panum sur la rétine, car le point M est situé dans l'horoptère de cette personne, il sera vu simple, en vision binoculaire, les images Md et Mg vont pouvoir fusionner. À l'inverse, le point N n'est pas dans l'horoptère, les images rétiniennes, Nd et Ng ne sont pas dans la même aire de Panum et ne seront donc pas fusionnées, le point N sera soit vu double soit par un seul œil (suppression de l'image N du deuxième œil).

Les zones de diplopie sont des zones dans lesquelles, il n'est pas possible d'obtenir une image binoculaire confortable ou d'obtenir une image binoculaire. Une diplopie est le phénomène de dédoublement de l'image, la personne ne peut pas converger ou diverger suffisamment pour fusionner correctement les images des deux yeux. Cela est généralement dû à un non-respect de la relation accommodation-convergence. Dans ce cas, le système visuel humain peut supprimer une ou des parties des images rétiniennes afin de joindre le plus grand nombre d'informations provenant des images rétiniennes gauches et droites. Cette suppression implique la perte, le masquage d'un certain nombre d'indices visuels. En d'autres termes, le cerveau va supprimer dans l'une des deux images les éléments qui devraient, mais ne peuvent pas être fusionnés avec les éléments de la deuxième image.

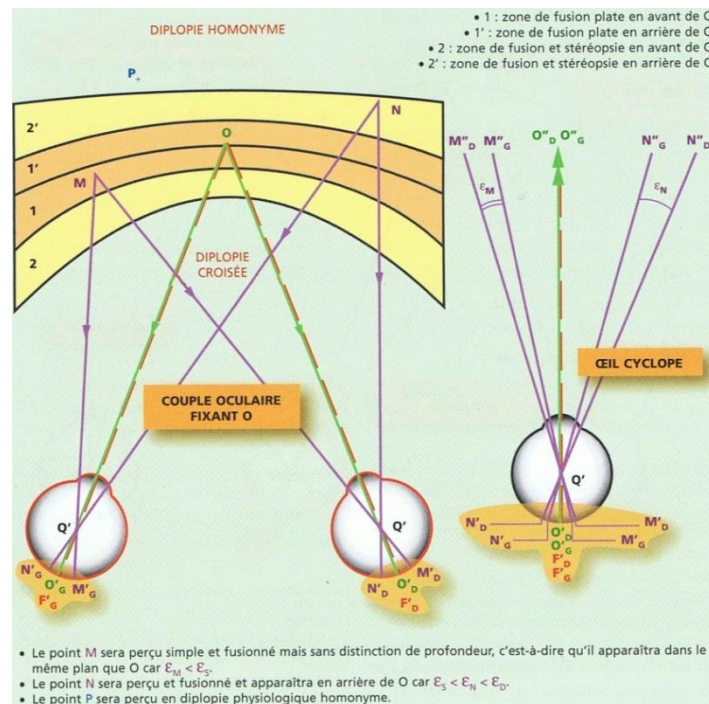


Figure 1.16 : Schéma de l'horoptère expérimental et œil cyclope (à droite) ; montrant les appariements de points entre les deux yeux¹⁸.

Le troisième et dernier niveau est la stéréoscopie : il s'agit du dernier niveau de la vision binoculaire. Ce niveau permet d'améliorer la perception de la perspective, des distances par l'analyse des disparités rétinienne entre les images rétinienne des deux yeux afin de calculer l'information de distance de ces points. La stéréoscopie se base sur l'analyse de la différence angulaire entre les images rétinienne pour calculer position dans l'espace d'un point, la différence entre $M_d - O_d$ et $M_g - O_g$ permet de calculer la position du point dans l'espace. Cette faculté visuelle dépend de chaque personne, elle peut être très développée comme inexistante. Il est possible pour la quantifier de mesurer une acuité dite stéréoscopique. La stéréoscopie nécessite le respect des deux précédents niveaux de la vision binoculaire avant de pouvoir être sollicitée.

¹⁸ Eric Beaubert, Franck Pariguet, et Stéphane Taboulot. *Manuel de l'opticien*. ISBN 2-224-02813-X. Editions Maloine, 2005

2. Analyse des besoins: aspects qualitatifs et quantitatifs

Dans ce chapitre, je vais présenter une analyse des besoins des personnes malvoyantes. Comme expliqué dans le chapitre 1, le système visuel est complexe, de plus il existe une grande diversité de trouble de la vision. Les besoins en fonction de ces troubles vont donc être différents. Certains troubles vont impacter la mobilité tandis que d'autres vont impacter la lecture. Dans la section 2.1, je vais commencer par rappeler les chiffres de la basse vision en France en Europe et dans le monde. Je vais également rappeler la classification des atteintes de la vision selon l'organisation mondiale de la santé. La section 2.2 sera consacrée à l'analyse des besoins des personnes malvoyantes. Dans la section 2.2.1, je vais commencer par présenter les données issues de plusieurs études réalisées par des organismes d'études statistiques ou des associations. Dans la section 2.2.2, je vais présenter les différents produits du marché destinés à apporter une aide pour des tâches où l'utilisation de la vision est importante. Dans la section 2.2.3, je vais présenter ma propre étude. J'ai réalisé cette étude afin d'apporter un certain nombre d'informations complémentaires. Dans la section 2.3, je vais, enfin, présenter l'immersion que j'ai réalisée au sein de la clinique ARAMAV, clinique spécialisée dans la réadaptation de personnes malvoyantes. Pour cela, je vais présenter cette clinique, son fonctionnement, puis une analyse des informations que j'ai obtenues de cette immersion.

2.1. Introduction

D'après l'organisation mondiale de la santé [126, 2, 125, 134], en 2012, il y a actuellement dans le Monde 285 millions de personnes ayant des troubles de leur vision. Il est recensé 246 millions de malvoyants.

D'après l'institut de la Vision en France, il y aurait 50 millions de malvoyants en Europe et aux États-Unis. Ce nombre pourrait également doubler dans les années à venir.

Aux Etats-Unis en 2000, il y avait environ 3,3 millions de malvoyants [44]. Ce chiffre est amené à croître de 70 % avant 2020.

En France, lors d'une étude de l'INSEE en 1999 [71, 165], il avait été recensé 1 700 000 déficients visuels dont 61 000 aveugles complets. Les malvoyants profonds étaient au nombre de 207 000 personnes. Ils présentaient une vision très limitée se limitant à la détection de silhouette, et de forme. Les malvoyants moyens étaient au nombre de 932 000 personnes. Ils présentent des difficultés pour la reconnaissance de forme, des visages, pour lire et écrire. Les malvoyants légers étaient au nombre de 550 000 personnes et présentent des déficiences légères.

Une partie de ces données a été actualisée en 2011 par l'Institut de la vision. Le Professeur José Alain Sahel et l'Institut de la Vision évoquaient 1 500 000 personnes concernées par la DMLA (dégénérescence liée à l'âge), 800 000 personnes concernées par le glaucome, entre 400 000 et 500 000 personnes concernées par la rétinopathie diabétique et 40 000 personnes concernées par la rétinite pigmentaire. Ces nouveaux chiffres donnent une estimation entre 2 740 000 à 2 840 000 malvoyants sans prendre en compte les autres maladies de la vue.

Selon l'OMS, il est possible de classer les déficiences visuelles, en 5 groupes :

- La déficience visuelle moyenne, qui se caractérise par un champ visuel supérieur à 20 degrés, et une acuité visuelle comprise entre 1/10 et 3/10.
- La déficience visuelle sévère, qui se caractérise par une acuité visuelle entre 1/20 et 1/10. La personne est alors capable de compter les doigts à 3 mètres.
- La déficience visuelle profonde, qui se caractérise par une acuité visuelle entre 1/50 et 1/20. La personne peut compter les doigts à 1 mètre.
- La déficience visuelle quasi-totale, qui se caractérise par une acuité inférieure à 1/50, un champ visuel réduit, et la perception de la lumière uniquement.
- La dernière classe correspond à une cécité totale, qui se caractérise par la non-perception y compris de la lumière.

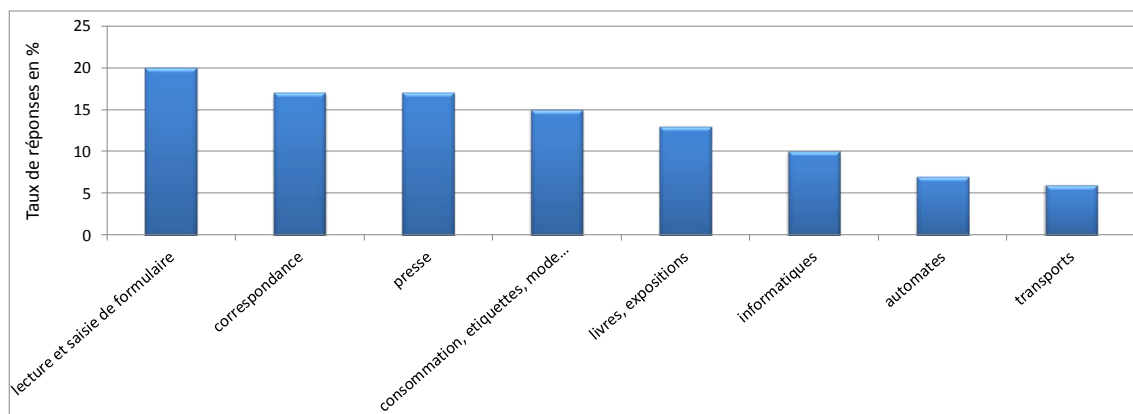
Il existe un certain nombre de déficiences visuelles pouvant atteindre plus ou moins sévèrement la vision du sujet atteint. Comme nous venons de le voir, l'OMS a défini différents niveaux de basse vision afin de classer les différents patients en fonction de leurs capacités visuelles résiduelles. Nous allons dans un premier temps (section 2.2.1), analyser les données issues d'une étude sur les usages et besoins des personnes malvoyantes. Par la suite, nous allons présenter différentes solutions actuellement disponibles sur le marché. Nous

allons ensuite analyser l'adéquation de ces solutions aux besoins (section 2.2.2) et réaliser une analyse plus approfondie de certains besoins par un questionnaire (section 2.2.3). Celui-ci a pour objectif de choisir les bonnes pistes de développement et de recherche, afin d'apporter le plus d'aide possible aux patients. Enfin, je vais présenter l'analyse issue de mon immersion au sein d'une clinique de réadaptation de personnes malvoyantes (section 2.3).

2.2. Analyse quantitative des besoins

2.2.1. Les enquêtes existantes

Nous allons commencer par analyser les besoins exprimés par les adhérents de l'association HandicapZéro¹⁹ lors de son enquête de 2005²⁰. Cette étude a été choisie, car elle regroupe l'avis de plus de 700 malvoyants Français, de plus il y est analysé un certain nombre de besoins. Il est important de noter que les besoins exprimés dans cette étude sont très simples et pourtant très importants pour la vie de tous les jours. De même, les aides techniques proposées afin de répondre à ces besoins sont très simples, cependant, elles ne sont pas forcément accessibles à tous.

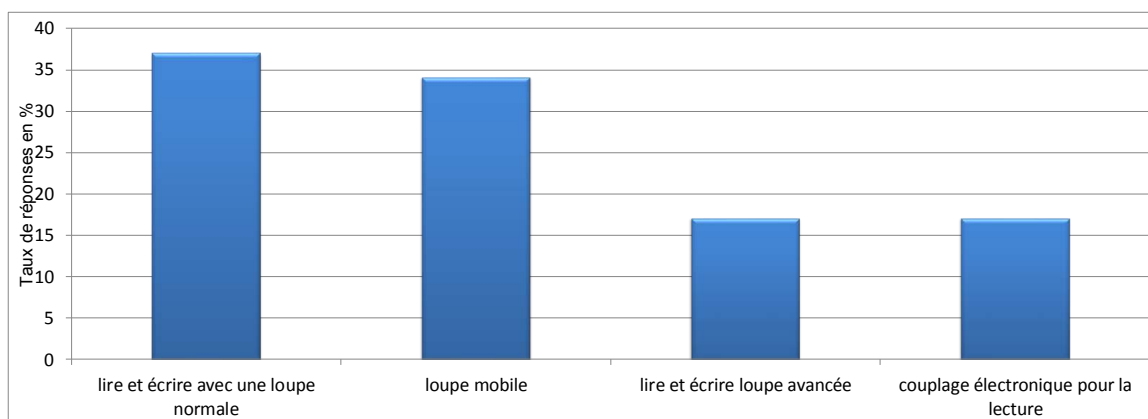


(a) Modes de lecture

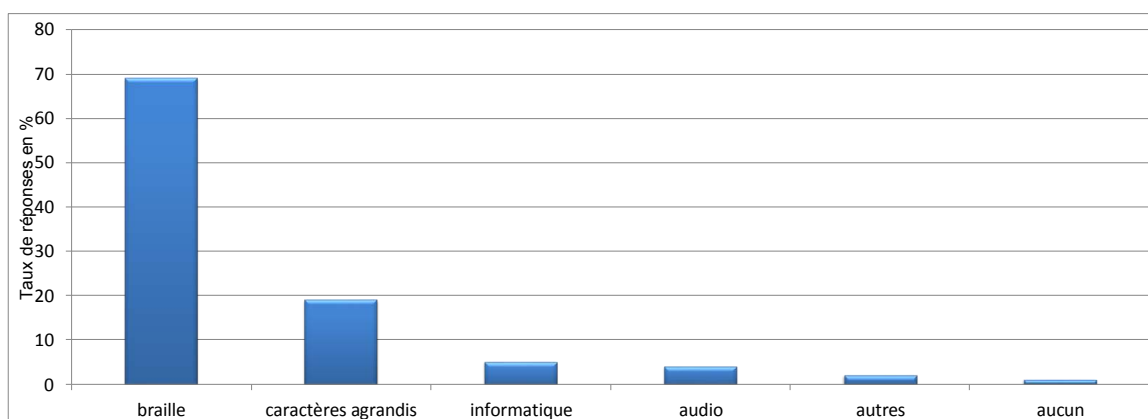
¹⁹ <http://www.handicapzero.org/>

²⁰

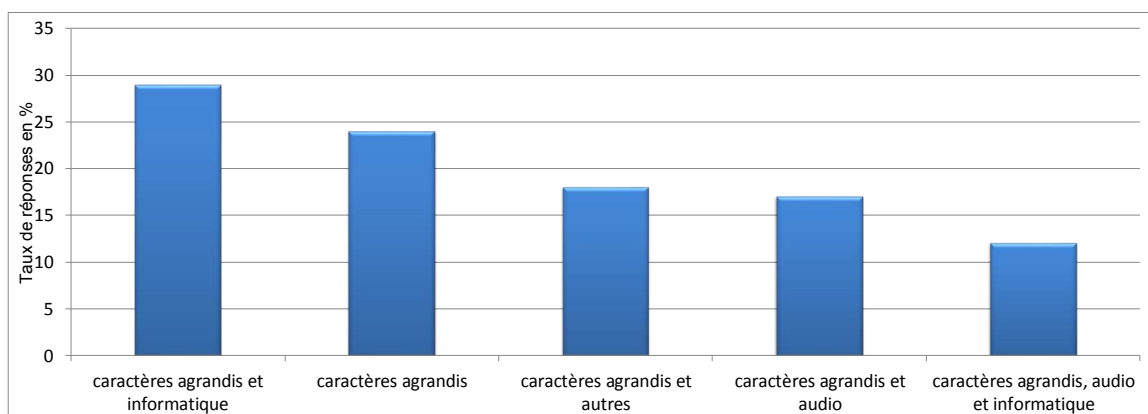
http://www.handicapzero.org/fileadmin/user_upload/contenu/enquete/telechargez_l_enquete_compleete_en_pdf.pdf. Site Handicapzéro



(b) Solutions techniques



(c) Mode de lecture des malvoyants et non-voyants



(d) Solutions techniques

Figure 2.1 Modes, de lecture, utilisés par les malvoyants et les solutions techniques, données compilées par Laster à partir de l'étude handicapzéro

Il est intéressant de voir que les malvoyants utilisent, cf. Figure 2.1, notamment les caractères agrandis pour lire. Cependant, il est souvent fait usage de fonctions complémentaires ou technologies complémentaires, cf. Figure 2.2 afin de réaliser la tâche de lecture plus aisément.

Il est en effet, souvent utilisé le son, le braille et/ou des méthodes informatiques couplant le son²¹, l'image^{22 23} et le braille²⁴ pour lire.

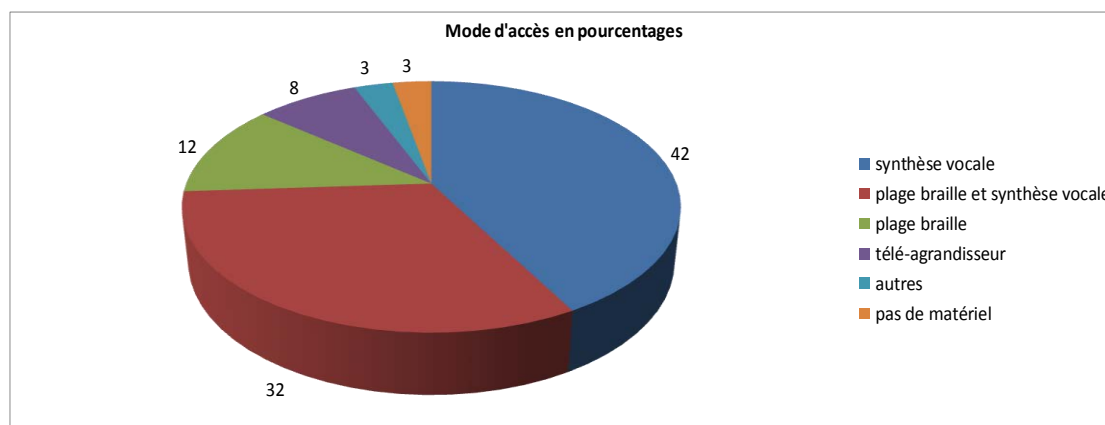


Figure 2.2 Mode d'accès à des contenus multimédia pour les personnes malvoyantes, en pourcentages, données compilées par Laster à partir de l'étude HandicapZéro

Il est intéressant de faire un lien entre les besoins et les équipements utilisés. L'utilisation des télé-agrandisseurs et des outils informatiques reste marginale. Il est couramment utilisé, l'audio (livres-audio, Jaws²⁵ pour lire une page web), le braille afin de percevoir un contenu (plage braille connectée à Jaws). En effet, la synthèse vocale se révèle intéressante pour utiliser un téléphone portable lorsque les télé-agrandisseurs, ou les outils informatiques sont moins appropriés.

Il est également intéressant pour une aide visuelle de pouvoir coupler plusieurs outils afin de pouvoir étendre son utilisation principale à d'autres fonctions telles que du contenu audio, visuel, et tactile.

Activité	20-59 ans	60-75 ans	76 et +	tous
Mobilité, déplacements	22.6 %	59.4 %	85.2 %	56 %
Tâches	19.8 %	47.5 %	76.3 %	47 %

²¹ <http://www.freedomscientific.com/products/fs/jaws-product-page.asp>

²² <http://www.aisquared.com/zoomtext>

²³ <http://www.yourdolphin.com/productdetail.asp?id=3>

²⁴ Il s'agit d'un outil réalisant la transcription de caractères de l'alphabet en signes brailles correspondants http://fr.wikipedia.org/wiki/Plage_braille

²⁵ <http://www.freedomscientific.com/products/fs/jaws-product-page.asp>

ménagères				
Alimentations, habillement	7.4 %	20.7 %	50.4 %	27.8 %
Communication, orientation	8.5 %	19.2 %	41 %	24.2 %

Figure 2.3 Besoins, en pourcents, des personnes **malvoyantes par tranches d'âge**, données du Ministère des affaires sociales et de la santé & INSEE - études et résultats juillet 2005 n° 416

On peut constater dans la Figure 2.3, que les personnes âgées expriment beaucoup plus de besoins que les personnes plus jeunes pour le même type de handicap. En effet, le handicap avec les autres problèmes de la personne vont lui faire perdre plus rapidement son autonomie. Il est donc nécessaire de prendre en compte, l'âge des personnes concernées, lors de la conception d'aide à la personne. Le patient doit pouvoir utiliser ce système sans effort avec une interface qui lui permet d'utiliser ce type de solution de manière simple, ergonomique et aisée.

D'une manière générale, indépendamment de l'âge, les patients expriment dans :

- 26.8 % des cas, un besoin, une nécessité d'utiliser une aide à la marche et pour les déplacements
- 13.5 % des cas, un besoin pour des aides optiques ou électroniques de type loupe
- 1.9 % des cas, un besoin par la canne blanche
- 0.5 % des cas, un besoin pour une interface d'ordinateur adaptée²⁶
- 0.3 % des cas, un besoin pour une solution d'assistance à la lecture soit de type logiciel informatique, soit une machine à lire

Beaucoup d'aides ont été développées pour la lecture et la vision de près. Cependant, on constate aussi un besoin important pour la mobilité, et les tâches quotidiennes, pour autant très peu de solutions sont actuellement disponibles sur le marché pour aider les malvoyants dans ces activités vitales.

L'étude qui a amené au développement d'outils interactifs, tel que VizWiz²⁷ sur des smartphones, a permis de mieux comprendre les besoins des malvoyants [31]. Cette étude a

²⁶ Il s'agit de modifier l'aspect habituel du logiciel pour le mettre en forme en mode bicolore, noir-blanc, et/ou agrandi, de manière à adapter l'interface aux besoins visuels de la personne

mis en avant certains besoins importants pour la reconnaissance d'objets, principalement des objets alimentaires, mais également beaucoup d'objets du quotidien. Cette étude a également mis en évidence les difficultés rencontrées par les malvoyants pour regarder la télévision, identifier et reconnaître les personnages, les détails. Pour autant, ces difficultés ne semblent pas ou peu influencer leur usage de la télévision par rapport aux personnes bien voyantes [197].

Notons également comme l'ont analysé J. Wobbrock et K. Shinohara [173], que les malvoyants et les personnes aveugles ne souhaitent pas que leur déficience puisse être vue quand ils utilisent des appareils adaptés à leur handicap. Il est ainsi fait mention d'un patient préférant utiliser un iPhone au lieu d'une plage braille adaptée, car celui-ci ne met pas son handicap à la vue de tous, et lui permet de bénéficier du même produit que tout le monde et des mêmes standards, comme le wifi. De plus, ce patient a indiqué que la plage braille coûte 6 fois plus cher que l'iPhone. Pour leur sécurité et celles des autres, les patients souhaitent informer les personnes de leur déficience : par exemple par le biais de la canne blanche. Enfin, ils souhaitent bénéficier d'aide de la part des autres personnes, mais sans excès.

Notons également que l'ergonomie des produits doit permettre aux patients une utilisation facile. Plusieurs études ont montré qu'il fallait faire participer les patients, lors du design et de la conception des aides [120, 156]. Une aide, pour être acceptée doit être fonctionnelle, mais également esthétique. Les patients expriment souvent des inquiétudes quant à l'impact du produit sur leurs relations avec les autres personnes.

2.2.2. Les produits du marché

Nous allons maintenant présenter la plupart des aides disponibles sur le marché. Cette synthèse ne sera pas exhaustive, car beaucoup de ces aides se ressemblent et proposent les mêmes fonctions. J'ai choisi de présenter uniquement les aides pour lesquelles nous disposons de suffisamment d'informations.

Les aides visuelles se classent en plusieurs catégories en fonction de leur poids, leur grossissement, leur taille, leurs fonctionnalités d'affichage, leur autonomie, leur portabilité et leur ergonomie. Il est également possible de classer les aides visuelles par type de dispositifs : loupe optique, loupe électronique, télé agrandisseur ou vidéo agrandisseur, machine à lire et logiciel pour ordinateur. Il existe également d'autres types de dispositifs d'aide que nous

²⁷ <http://vizwiz.org/>

n'analyserons pas ici, car ils sont moins proches des fonctionnalités ciblées dans notre projet et n'utilisent pas de traitement d'image : canne blanche, canne blanche électronique, GPS, système d'assistance personnel, guide vocal et livre audio.

Il existe sur le marché tout un ensemble d'aides non-visuelles pour:

- La mobilité : la canne blanche, le chien-guide, le GPS vocal, des capteurs ultrason pour la perception de l'information de distance, des aides au franchissement de feux piétons,...
- La lecture : machine à lire, plage braille ...
- La vie quotidienne : des lecteurs de couleurs, livres vocaux ainsi que des appareils pour la vie de tous les jours, tels que les balances parlantes, les réveils parlants ...

La liste des dispositifs présentés ci-après est uniquement donnée à but indicatif. L'objectif est avant tout de présenter les fonctionnalités offertes par ces dispositifs, leurs inconvénients, leurs contraintes (poids, coût) et leurs limites. Cette liste n'a pas vocation à être exhaustive. Cette étude couplée à celle sur des besoins des personnes malvoyantes, va nous permettre de mettre en avant, de démontrer l'intérêt des solutions complémentaires à celles actuellement disponibles.

2.2.2.1. Les loupes optiques

Les loupes optiques permettent d'agrandir les images ou les objets observés, avec des grossissements variables d'une loupe à l'autre, et avec ou sans éclairage. Les grossissements disponibles sont limités à 5 ou 8 fois la taille de l'objet original.

2.2.2.2. Les loupes électroniques

Les loupes électroniques, cf. Figure 2.4, permettent de remplir les mêmes fonctions que les loupes optiques, mais avec des grossissements plus importants et des fonctions telles que la modification des couleurs, ou un arrêt sur image.

Produit	Pebble	Amigo	LEP 430	Mano
Poids	210g	600g	221g	139g
Dimensions en mm	78 X 111 X 35	95 X 165 X 40	146 X 178 X 23	89 X 74 X 21

Energie	Piles	Batterie	Batterie	Batterie
Autonomie	2 heures	1 heure 30	5 heures	2 heures 30
Modes	Couleurs, fausses couleurs, noir et blanc	Couleurs, fausses couleurs, noir et blanc	Couleurs, fausses couleurs, noir et blanc	Couleurs, fausses couleurs, noir et blanc
Agrandissement	Jusqu'à 10X	Jusqu'à 14X	4X à 22X	1X à 20X
Mode photo	Oui	Oui	Oui	Oui
Prix	650 €	2290 €	990 €	1495 €

Figure 2.4 Tableau comparatif des loupes électroniques, données issues des fiches techniques et catalogues des distributeurs en France^{28, 29, 30, 31, 32}.

2.2.2.3. Les machines à lire

Les machines à lire, cf. Figure 2.5, permettent d'écouter un document, ou un livre sans avoir à le lire. Ces dispositifs associent une numérisation, une reconnaissance des caractères et un processus de synthèse vocale pour la lecture du document.

Produit	ClearReader+	Lector	Narratis
Poids	2,46 Kg	7,5 Kg	7,5 Kg
Dimensions en mm	230 X 240 X 100	470 X 275 X 135	470 X 275 X 100
Energie	Batterie et secteur	Secteur	Secteur
Autonomie	5 heures	Non indiqué	Non indiqué
Son	Haut-parleurs, écouteurs	Haut-parleurs, Ecouteurs Lecteur audio daisy	Haut-parleurs, Ecouteurs Lecteur audio daisy

²⁸ <http://www.avh.asso.fr/>

²⁹ <http://www.axos.eu/>

³⁰ <http://www.celyatis.com/>

³¹ <http://www.etexfrance.fr/>

³² <http://www.visiole.fr/>

Reconnaissance	OCR	Scanner OCR	Scanner OCR
Prix	3190 €	3140 €	2550 €

Figure 2.5 Tableau comparatif des machines à lire, données issues des fiches techniques et catalogues des distributeurs en France³³.

2.2.2.4. Les télé-agrandisseurs

Les télé-agrandisseurs ou vidéo agrandisseurs, cf. Figure 2.6, sont des dispositifs d'aide aux malvoyants bien connus. Ce type de dispositif est généralement lourd et assez coûteux. Il permet d'atteindre des grossissements très élevés et des modifications de couleurs au moment de l'affichage.

Produit	Zip	Naratello	Phénix +	Topolino smart	Mezzo
Poids	8 Kg	Non indiqué	17 Kg	2,3 Kg	4,7 Kg
Dimensions en mm	400 X 430 X 530	930 X 810 X 560	510 X 600 X 580	280 X 445 X 450	320 X 450 X 460
Autonomie	Secteur	Secteur	Secteur	2 heures 30	3 heures 30
Agrandissement	Non indiqué	2,3X à 28X	2,5X à 95X	1,9X à 85X	2,8X à 62X
Options	Pliant	Lecteur vocal	Non indiqué	Pliant	Piant
Prix	2975 €	3950 €	3995 €	4490 €	3200 €

Figure 2.6 Tableau comparatif des télé-agrandisseurs, données issues des fiches techniques et catalogues des distributeurs France.

2.2.2.5. Les solutions informatiques

Les logiciels, cf. Figure 2.7 pour personnes non-voyantes ou mal voyantes sont de plusieurs types, ils varient en fonction des fonctionnalités proposées. Ils permettent, en effet, d'obtenir des apports d'information sous forme d'agrandissement d'écran, de lecteur d'écran, de lecteur et commande vocale.

³³ <http://braille-vocal.cecjaa.com/clearreader-machine-a-lire-avec-garantie-2-ans-c2048-159.php>

Produit	Zoom text	Acapella	Supernova suite	Jaws
Agrandissement écran	1X à 36X	Non indiqué	1,2X à 60X	Non indiqué
Lecture écran	Oui	Oui	Oui	Oui
Prix	760 €	595 €	1490 €	799 €

Figure 2.7 Tableau comparatif des logiciels informatiques, données issues des fiches techniques et catalogues des distributeurs en France

Au vu de la section 2.2.1 et de cette section, nous pouvons constater que les patients ont des beaucoup de besoins auxquels doivent répondre les dispositifs d'aides, alors que les dispositifs actuels sont spécialisés pour l'aide à la lecture. Au regard du taux d'équipement et des besoins exprimés par les patients, je peux en déduire que les solutions actuellement proposées ne sont pas en adéquation avec tous les besoins et toutes les situations. Ces dispositifs répondent à un certain nombre de situations, cependant, il existe une différence notable entre les besoins exprimés et les usages que l'on peut faire de ces aides. Cela peut s'expliquer par le prix, les fonctions proposées et la portabilité des solutions. Un certain nombre de solutions performantes ne sont peu ou pas transportables. Les solutions mobiles actuelles sont, quant à elles, généralement peu performantes. Il est nécessaire de développer un nouveau type de solution offrant à la fois les performances, la mobilité, et la portabilité. Au vu de ce constat, il apparaît utile d'analyser plus en détail les besoins des personnes malvoyantes afin d'approfondir cette analyse. Cette analyse devra nous permettre de mieux définir les pistes à explorer pour développer cette nouvelle aide.

2.2.3. Notre étude menée à partir d'un questionnaire

Afin d'approfondir les données dont nous avons connaissance concernant les besoins des malvoyants, j'ai établi un questionnaire, cf. Annexe 1, destiné à évaluer les besoins des déficients visuels. Plus précisément, j'ai évalué les besoins en vision de près (par exemple pour la lecture), pour l'utilisation d'écrans d'ordinateur, pour regarder la télévision, pour la mobilité et pour l'utilisation de systèmes de communication, tels que les téléphones ou les tablettes numériques.

Puis j'ai posé les mêmes questions à deux groupes de malvoyants provenant pour le premier de L'Institut ARAMAV³⁴ et pour le deuxième de l'association Rétina France³⁵ – Vision & recherche³⁶. Ces malvoyants qui ont participé à notre étude, ont été choisis afin d'avoir suffisamment de participants représentatifs dans chacune des classes de malvoyants (cf. chapitre 1 section 1.1.5).

Par la suite, le groupe 1 désigne les malvoyants interrogés dans le cadre de l'ARAMAV. De même, le groupe 2 désigne les malvoyants interrogés dans le cadre de Rétina France. Lorsque les données issues des deux groupes de malvoyants ont été fusionnées, il est indiqué « global ».

2.2.3.1. Catégories des malvoyants interviewés

J'ai analysé chaque paramètre caractéristique de chaque malvoyant et chaque réponse. Lorsqu'on analyse l'âge des malvoyants, cf. Figure 2.8, il n'y a pas parmi les deux groupes de différences statistiques, j'ai donc considéré les deux groupes comme provenant de la même population de malvoyants.

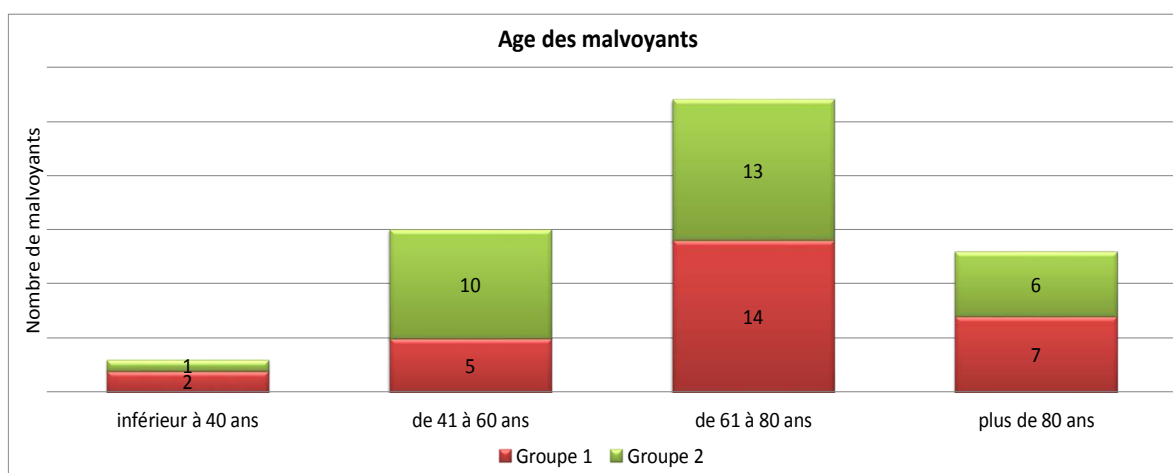


Figure 2.8 Répartition de l'âge des malvoyants ayant répondu au questionnaire

Lorsque l'on compare la répartition des malvoyants en fonction de leur âge entre les deux groupes, le test du khi deux permet de considérer comme un seul et même groupe ces deux populations statistiques, la probabilité pour que l'hypothèse de deux groupes différents est bien inférieure à 0,05.

³⁴ <http://www.aramav.com/pilot.html>

³⁵ <http://www.retina.fr/>

³⁶ <http://www.retina.fr/spip.php?article78>

Lorsque l'on regarde les professions, cf. Figure 2.9, des malvoyants interrogés, il n'y a pas non plus de différences. Il y a bien quelques petites différences, par exemple entre le nombre d'étudiants et le nombre de travailleurs à temps partiels, mais ces différences sont mineures. Le premier groupe était composé de 28 malvoyants, le deuxième groupe était composé de 30 malvoyants. Nous avons donc interrogé 58 malvoyants au total.

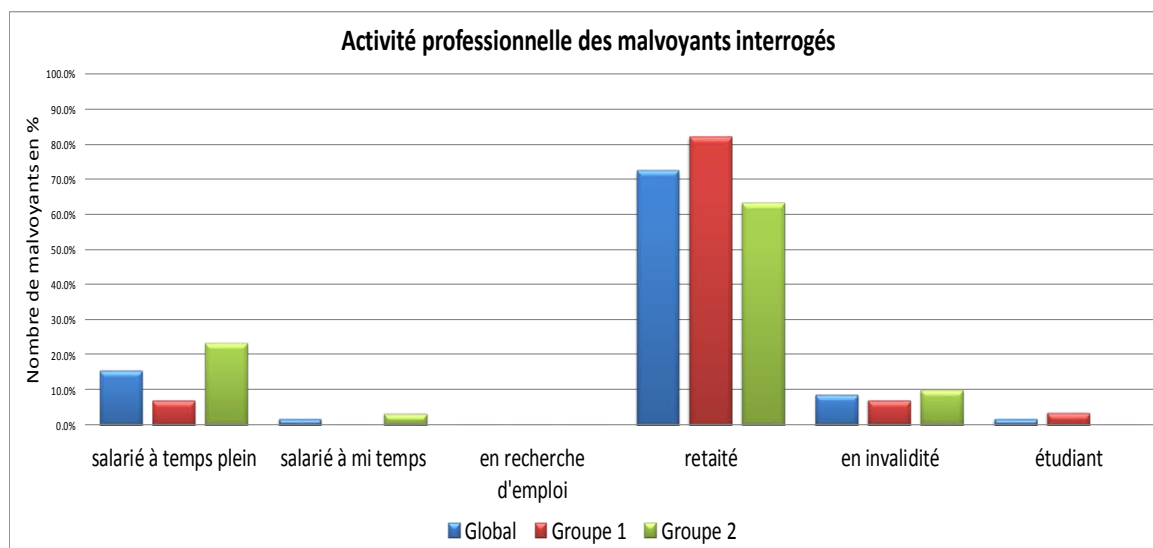


Figure 2.9 Catégorie socio-professionnelles des malvoyants en pourcentage par rapport à chaque groupe.

Lorsque l'on compare la répartition des malvoyants en fonction de leurs catégorie socio-professionnelle entre les deux groupes, le test du khi deux permet de considérer comme un seul et même groupe ces deux populations statistiques, la probabilité pour que l'hypothèse de deux groupes différents est bien inférieure à 0,05.

Certains malvoyants ont expliqué avoir pris leur retraite lors de l'apparition de leur maladie ou avoir été placé en invalidité lors de l'apparition de celle-ci.

Une des informations les plus importantes lorsque l'on parle de malvoyance est l'acuité visuelle. Nous pouvons voir sur la Figure 2.10 la répartition des acuités en vision de loin, exprimées en dixièmes, de chaque œil pour chacun des malvoyants. Pour un certain nombre de malvoyants, l'acuité à l'œil gauche est différente de celle de l'œil droit. De nouveau, au vu de ce critère, les deux groupes sont sensiblement identiques.

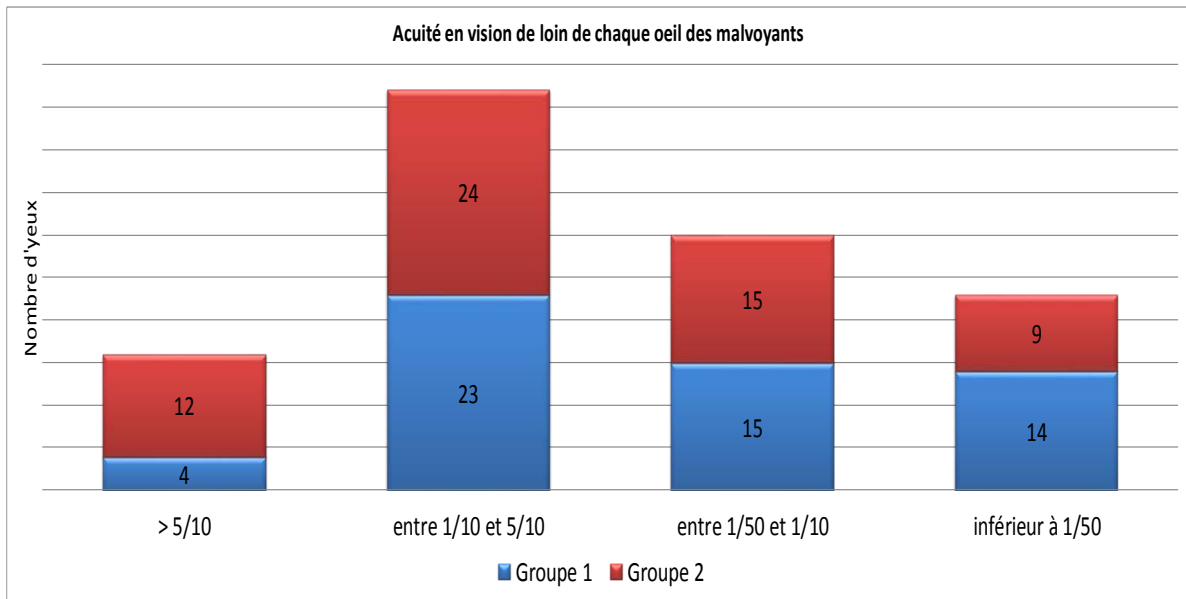


Figure 2.10 Répartition des acuités monoculaires en vision de loin.

Lorsque l'on compare la répartition des malvoyants en fonction de leurs acuités monoculaires entre les deux groupes, le test du khi deux permet de même de considérer comme un seul et même groupe ces deux populations statistiques.

Le deuxième paramètre important est le type d'atteinte du champ visuel. Nous avons pour cela classé les malvoyants en fonction de leur atteinte : scotome, tubulaire, excentré ou autre. Les atteintes de type « excentrement » et « autre » peuvent être cumulées avec une autre atteinte sur la Figure 2.11.

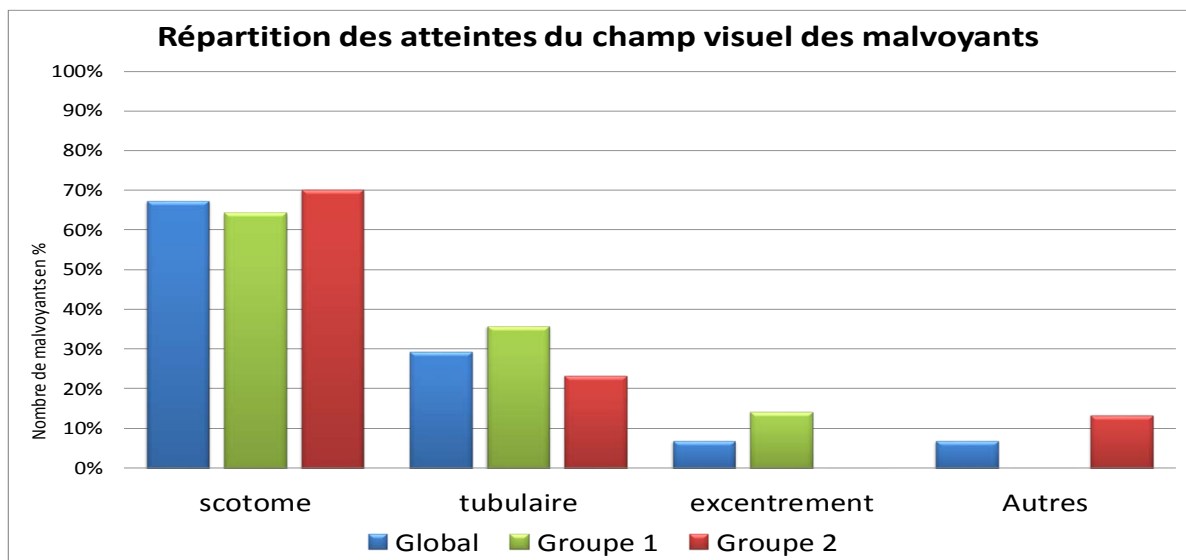


Figure 2.11 Répartition des malvoyants en fonction de l'atteinte du champ visuel.

Lorsque l'on compare la répartition des malvoyants en fonction de leur atteinte du champ visuel entre les deux groupes, le test du khi deux permet de même de considérer comme un seul et même groupe ces deux populations statistiques.

Pour rappel, le scotome concerne la DMLA, il s'agit de la destruction de la partie centrale du champ visuel correspondant à la fovéa et parfois à la macula. L'atteinte tubulaire, concerne la rétinite pigmentaire, il s'agit d'un rétrécissement du champ visuel de l'extérieur vers l'intérieur de celui-ci. Dans ce cas, le malvoyant voit comme à travers un trou de serrure. L'excentrement correspond au cas où la vision tubulaire ou le scotome ne sont pas axés sur le centre du champ visuel, mais décalés par rapport à celui-ci. Le cas de l'atteinte autre correspond à des pathologies n'affectant pas le champ visuel tel que l'albinisme.

2.2.3.2. Besoins exprimés par ces malvoyants

Nous allons maintenant nous pencher sur les besoins des malvoyants. Nous les avons notamment interrogés sur leurs connaissances des systèmes d'aide visuelle, cf. Figure 2.12, de leur usage, de leur fréquence d'usage, et de leur intention d'acquérir de tels systèmes, cf. Figure 2.13.

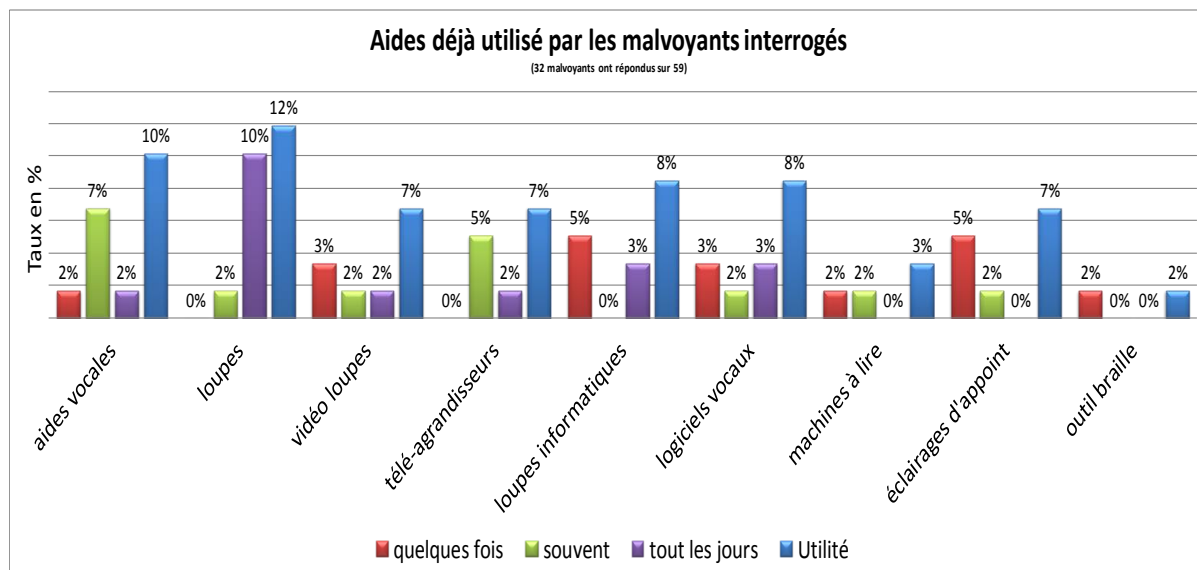


Figure 2.12 Taux de réponses pour chaque type d'aide visuelle et l'utilité de ces aides vis-à-vis de certains usages

On constate que les malvoyants utilisent peu les aides visuelles disponibles sur le marché. On peut aussi noter que les malvoyants âgés de plus de 75 ans ont des difficultés à maîtriser un certain nombre de ces aides et ne les utilisent pas ou très peu. Les malvoyants interrogés étant

en réadaptation, ils utilisent un peu moins les aides visuelles que des malvoyants après la réadaptation.

La Figure 2.12 permet également de constater que le taux d'équipement des malvoyants est bien de 10 à 15 % comme le montre l'étude de l'INSEE. Nous avons également demandé aux malvoyants s'ils avaient l'intention d'acheter certaines aides actuellement disponibles sur le marché.

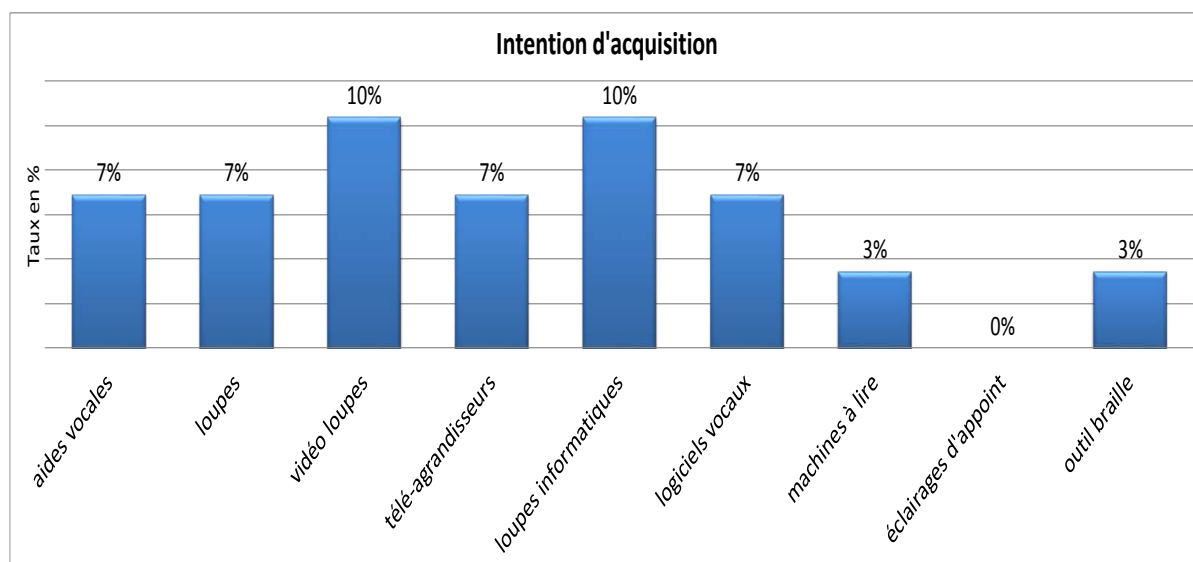


Figure 2.13 Taux d'intention d'acquisition d'une aide visuelle.

Peu de malvoyants ont l'intention d'acheter des aides visuelles, cf. Figure 2.13. Les malvoyants de plus de 80 ans expliquent ce choix par leur âge et le fait que les aides ne sont pas toujours faciles à prendre en main. Les autres malvoyants expliquent qu'ils utilisent seulement quelques aides visuelles auxquelles ils s'adaptent pour réaliser certaines des tâches dont ils ont besoin. Les malvoyants ont indiqué que le prix, la facilité d'utilisation et l'apport des aides constituent trois autres critères d'importance pour eux.

Nous constatons que les aides actuelles ne sont que partiellement en adéquation avec les besoins des malvoyants comme l'avait montré l'étude dont la HID 1999 [71].

Nous allons maintenant étudier l'impact des troubles visuels sur la vie quotidienne des malvoyants que nous avons interrogés. Cette étude va nous permettre de mieux comprendre les besoins des malvoyants et ainsi de leur proposer des solutions plus adaptées.

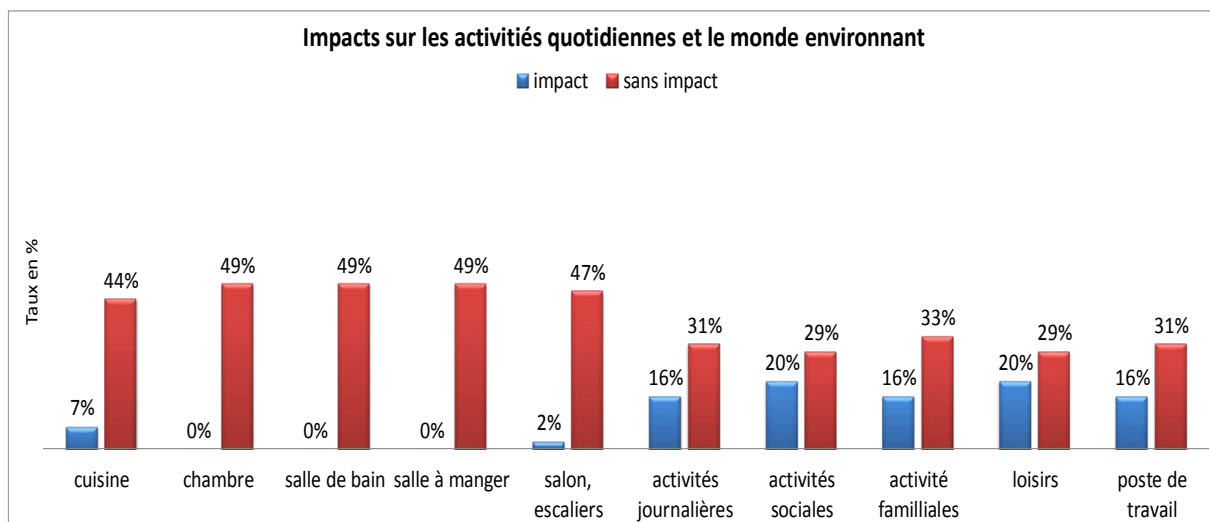


Figure 2.14 Impacts dus à la déficience visuelle.

Les pourcentages cumulés indiqués sur la Figure 2.14 ne sont pas égaux à 100 % car tous les malvoyants n'ont pas répondu à cette partie du questionnaire pour différentes raisons, dont le manque de temps. Certains malvoyants étant des patients de l'ARAMAV, ils ont répondu au questionnaire sur leur temps libre en donnant la priorité à leur programme au sein de la clinique.

Certains malvoyants interrogés ont indiqué qu'ils ont adapté leurs activités en fonction de leurs atteintes visuelles. D'autres ont maintenu leurs activités malgré leurs troubles visuels en utilisant quand cela était possible une aide adaptée. La plupart des malvoyants interrogés cherchent en général le maintien de leur activité professionnelle en fonction des solutions trouvées pour palier leurs déficiences. D'autres ont dû arrêter leur activité professionnelle. Dans la plupart des cas la déficience visuelle impacte leur vie et leur lien social (au contact de leur famille et/ou de leurs amis). Ces malvoyants d'une manière générale expriment le souhait d'avoir à disposition des solutions plus adaptées à leurs besoins.

Nous avons ensuite demandé à ces malvoyants de répondre à d'autres questions afin de déterminer dans quelles situations, ils se sentent en autonomie, ou pas. Le terme autonomie est employé ici pour indiquer une situation où le malvoyant peut évoluer, avec ou sans aide, avec ou sans personne à proximité et pour laquelle, le malvoyant est en confiance et sait comment réagir. Cette notion est importante, car certains malvoyants âgés ne sortent jamais seuls, de même certains malvoyants ne sortent jamais sans leur canne blanche ou leur chien-guide. Le fait que les malvoyants soient en confiance pour réaliser telle ou telle action est très important, car ils peuvent la renouveler par eux même.

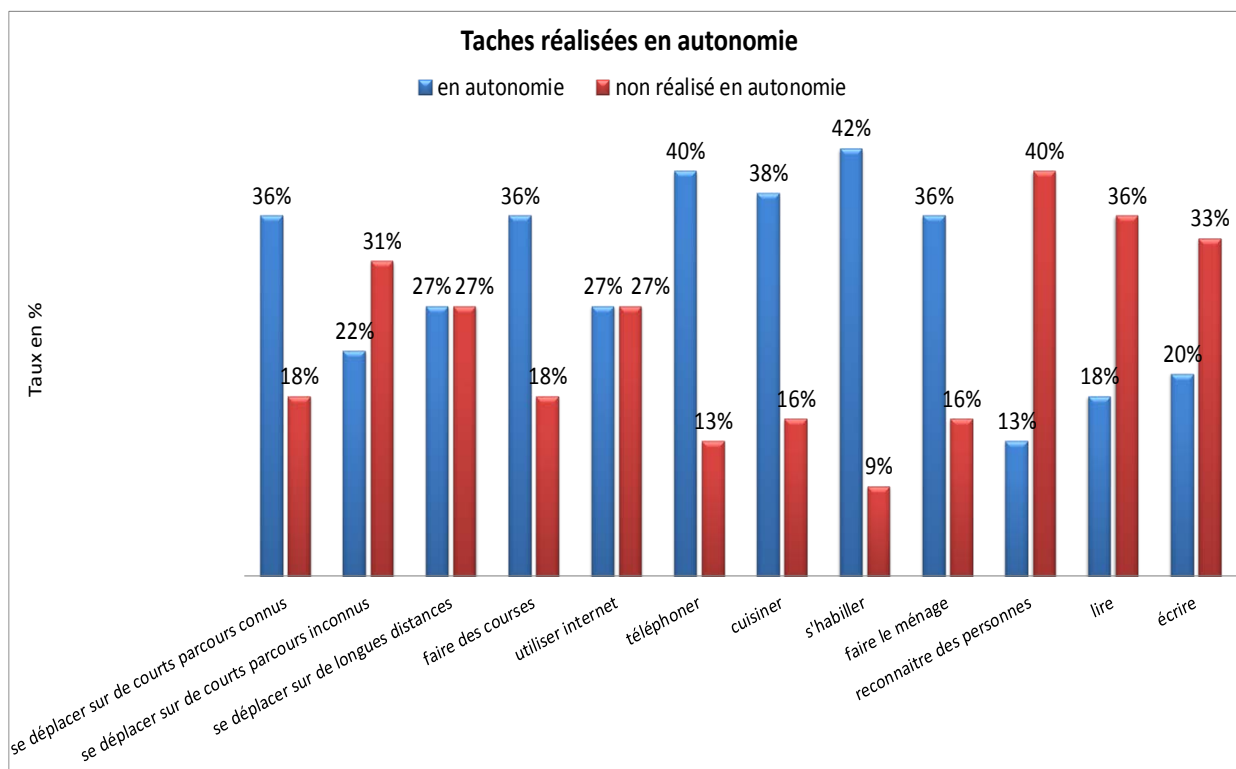


Figure 2.15 Taux de malvoyants estimant pouvoir réaliser certaines actions en autonomie.

Les pourcentages cumulés indiqués sur la Figure 2.15 ne sont pas égaux à 100 % car tous les malvoyants n'ont pas répondu à cette question. Les 2/3 des malvoyants ayant répondu estiment pouvoir se déplacer en autonomie sur des parcours courts connus. Pour les tâches de lecture, écriture, internet ou reconnaissance des visages, les malvoyants ne se sentent pas aptes à réaliser ces tâches en autonomie. Pour les autres tâches proposées, dans l'ensemble, les malvoyants se sentent autonomes, cependant la réponse de chaque malvoyant dépend de son âge, de son acuité, de sa pathologie visuelle et de ses besoins.

On peut cependant noter que les tâches importantes sont la mobilité sur des parcours inconnus, la lecture, l'écriture et la reconnaissance de personnes. Il semble intéressant pour les malvoyants de disposer d'une aide pouvant à la fois être utilisée en mobilité et pour lire, réaliser un travail en vision de près, et reconnaître des visages.

Afin de compléter cette étude, nous avons également posé une série de questions concernant les fonctionnalités existantes ou non, cf. Figure 2.16, dans les systèmes d'aide actuels et dans un système d'aide idéal. L'objectif de ces questions est de compléter notre compréhension des besoins par une analyse des usages souhaités.

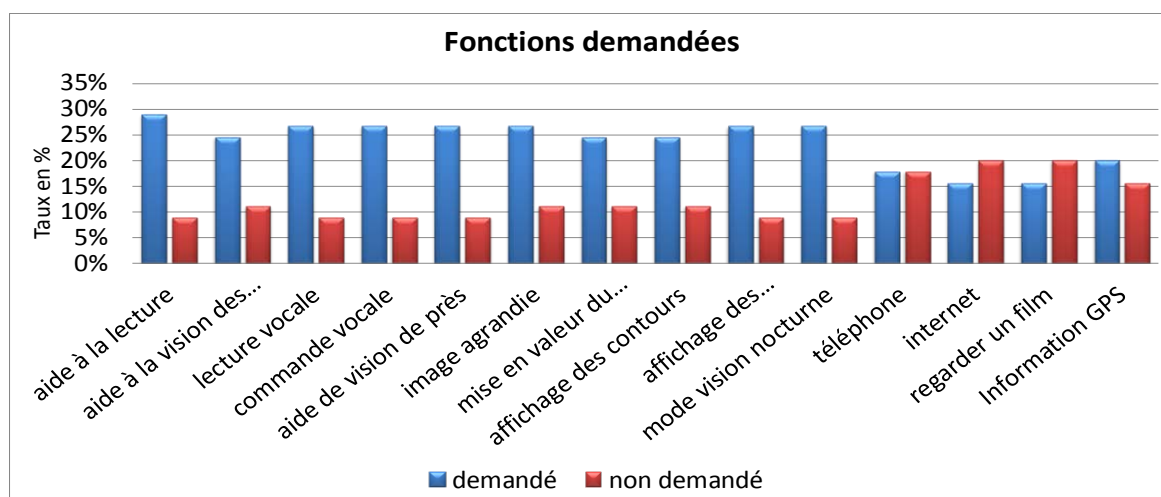


Figure 2.16 Fonctionnalités demandées

Les fonctions les plus demandées (réponse supérieure ou égale à 20 %) sont les fonctions vocales, de lecture (OCR et synthèse vocale), et une aide de vision de près (agrandissement, modification des couleurs, rehaussement). Ces fonctions correspondent à celles proposées par la plupart des télé-agrandisseurs et loupes électroniques, mais les malvoyants considèrent que ces systèmes sont difficilement utilisables en mobilité et transportables. Les fonctions de mise en exergue du texte, ou des contours, et de vision nocturne, ne sont pas disponibles actuellement. Pour autant, les malvoyants considèrent que ces fonctions pourraient être utiles pour eux.

2.2.3.3. Aspects ergonomiques

Nous allons dans un premier temps analyser les besoins en autonomie d'un système d'aide aux malvoyants. Les malvoyants ont indiqué qu'une autonomie allant jusqu'à 4 heures est acceptable, cf. Figure 2.17. Les malvoyants étant presque tous retraités, ils souhaiteraient pouvoir utiliser ce type d'aide, en mobilité, quelques heures par jour. Les malvoyants actifs ont expliqué que si le système est rechargeable, avec une batterie externe par exemple, cela peut répondre à leurs besoins.

Nous avons ensuite posé la question du poids. Nous avons notamment demandé aux malvoyants interrogés quel est le poids le plus raisonnable pour un dispositif porté sur tête, comme des lunettes de Galilée (système optique grossissant), ou pour tout autre dispositif sur tête, ainsi que pour un appareil électronique porté à la ceinture, comme une loupe électronique dans un étui à la ceinture.

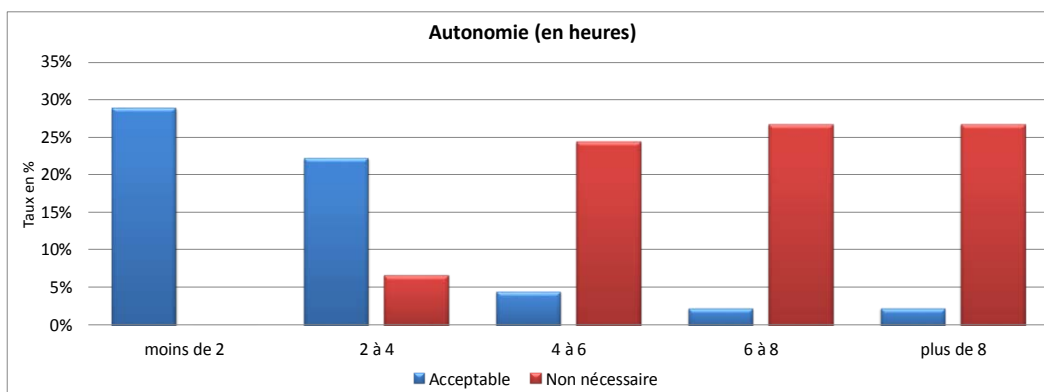


Figure 2.17 Autonomie acceptable pour une aide visuelle

Nous allons maintenant analyser les réponses concernant le poids du boîtier à la ceinture. Pour la majorité des malvoyants interrogés, le système peut peser jusqu'à 499 grammes, idéalement moins de 300 grammes, cf. Figure 2.18.

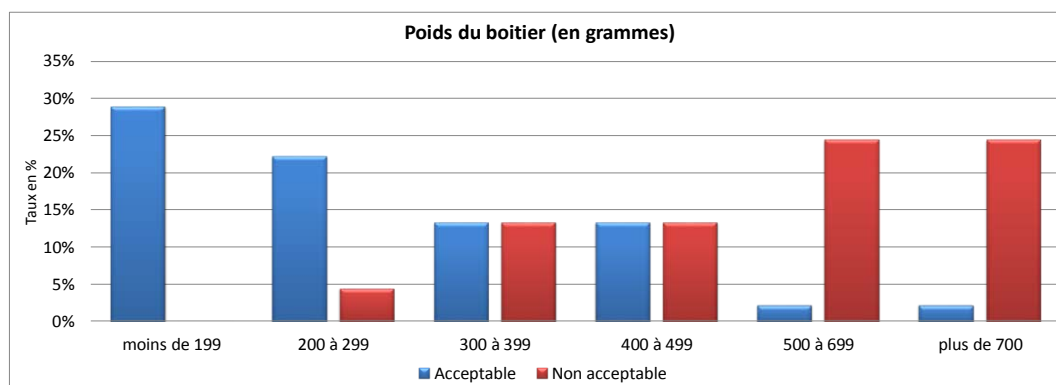


Figure 2.18 Poids du boîtier acceptable par les malvoyants pour une aide visuelle.

Finalement, les malvoyants interrogés ont indiqué quel pourrait être le poids acceptable pour un dispositif porté sur tête, cf. Figure 2.19. Pour la majorité des malvoyants interrogés, une paire de lunettes doit peser moins de 120 grammes, idéalement moins de 99 grammes.

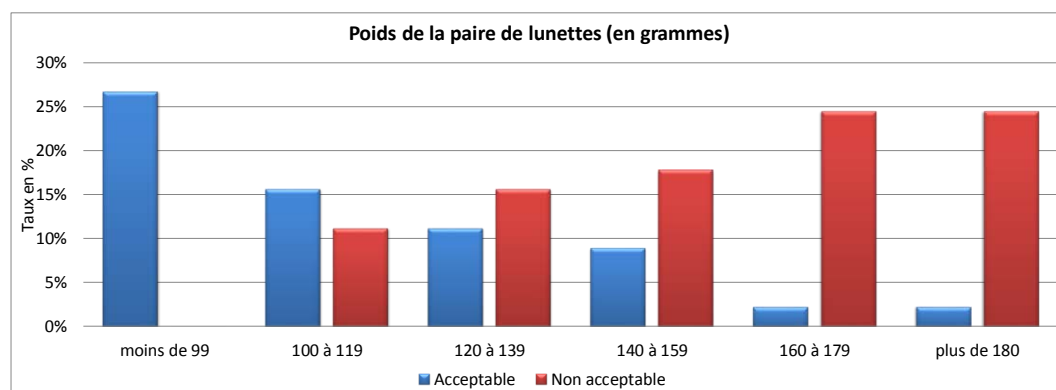


Figure 2.19 Poids acceptable pour une paire de lunettes d'aide visuelle.

2.2.4. Conclusion

J'ai présenté dans cette partie une analyse des besoins des malvoyants, leurs usages des solutions actuelles. Les malvoyants ont des besoins pour la lecture, l'écriture, la vision de près, ainsi que des besoins pour la mobilité, pour toutes les tâches de la vie quotidienne.

La réponse actuelle des dispositifs du marché est très concentrée sur ce premier besoin, la vision de près, comme le montre la deuxième partie. En effet, il existe beaucoup d'aides pour la lecture telle que les loupes optiques, loupes électroniques, télé-agrandisseurs, logiciel d'agrandissement et solutions de lecture de documents. Ces dispositifs au vu de leur faible taux d'utilisation semblent ne répondre que partiellement aux besoins auxquels ils sont sensés répondre. De plus, les solutions du marché ne répondent pas à tous les besoins des malvoyants. Il n'existe pas de solution d'aide visuelle pour la mobilité par exemple. De plus d'une situation à l'autre, les malvoyants doivent utiliser plusieurs aides pour réaliser le même type de tâche, par exemple utiliser une loupe pour lire l'affichage des horaires de bus et un télé-agrandisseur pour lire un courrier. Cela nécessite d'acheter et maîtriser un certain nombre d'appareils.

Nous avons, dans le but de mieux comprendre les besoins des malvoyants, réalisé notre propre questionnaire. Cela nous a permis d'une part de confirmer le faible taux d'équipement des malvoyants et d'autre part d'échanger avec eux afin de comprendre pourquoi ce taux est si faible et pourquoi ils utilisent si peu les dispositifs du marché. Cette étude est très instructive, car elle met en évidence un certain nombre de besoins, beaucoup plus importants que ceux analysés jusqu'ici. Les malvoyants ont exprimé d'une part, des besoins importants en vision intermédiaire et pour la mobilité et d'autre part, un besoin pour des solutions plus mobiles et plus polyvalentes que les solutions proposées actuellement. Ces informations nous ont permis de mieux comprendre les besoins de ces malvoyants, d'identifier les réponses que l'on pourrait apporter à ces besoins et ainsi définir de nouvelles recherches, en vue du développement de nouveaux dispositifs d'assistance visuelle.

2.3. Immersion

2.3.1. Contexte de l'immersion

Dans cette partie, je vais présenter une synthèse des informations collectées lors d'échanges avec les professionnels de la réadaptation et de la rééducation, ainsi que lors de discussions avec les patients. Cette étude a été réalisée dans le but de mieux comprendre les besoins des malvoyants, et comment notre travail peut s'intégrer en collaboration avec les professionnels, dans le parcours de soins, suivi par un patient. Ce positionnement est nécessaire pour apporter le plus d'aide possible à un patient tout en prenant en compte les éventuels inconvénients qu'une telle aide peut comporter et en limiter les effets autant que possible. Je vais tout d'abord réaliser une présentation de la clinique de l'ARAMAV au sein de laquelle j'ai réalisé une immersion de quelques jours au début de ma thèse. Je vais ensuite présenter les unités de réadaptation de la clinique puis les informations acquises auprès des professionnels et des patients, dans le but de comprendre au plus juste les besoins visuels des patients.

Cette phase d'immersion a été nécessaire car il est primordial de comprendre, d'appréhender de manière correcte les besoins des patients avant de réaliser des travaux de recherche et développements dans ce domaine. Il existe en effet, à ma connaissance peu ou pas de publications scientifiques présentant ces besoins. Il existe seulement un certain nombre de conférences nationales où il est présenté les nouvelles méthodologies utilisées par les professionnels de la réadaptation. On peut néanmoins citer les travaux de thèse de Guillaume Tatur [179] réalisés au sein de l'ARAMAV. Il existe en revanche un certain nombre de publications permettant d'analyser les performances de tel ou tel dispositif, ces publications seront évoquées dans le chapitre 3 (relatif aux développements algorithmiques).

En dehors de quelques références citées dans ce chapitre, les informations mentionnées proviennent des documentations de l'ARAMAV ou sont issues des discussions que j'ai menées avec les professionnels de la clinique.

2.3.1.1. Présentation de l'Institut ARAMAV

Dans le cadre des travaux présentés dans cette thèse, j'ai réalisé un partenariat avec l'Institut ARAMAV³⁷ à Nîmes. J'ai réalisé dans un premier temps, une phase d'immersion au sein de cette clinique dans le but de comprendre le fonctionnement de celle-ci, de voir comment il est possible de réadapter des personnes malvoyantes, et de comprendre quels sont les éléments visuels, ou non-visuels, utilisés par les patients.

La clinique du Belvédère est un institut de rééducation et de réadaptation fonctionnelle pour personnes aveugles ou malvoyantes. L'expression « réadaptation fonctionnelle » correspond ici à l'exploitation optimale de la vision résiduelle et des sens compensatoires, pour permettre le déplacement et les tâches de la vie quotidienne. Elle est gérée par l'Association pour la Réinsertion des Aveugles et Malvoyants (association loi de 1901 à but non-lucratif) dit AGE ARAMAV. La clinique a été créée en 1989.

L'objectif de cette clinique est de proposer des séjours de rééducation individualisés dans le cadre d'une prise en charge globale de la personne en fonction de sa prescription médicale et d'indications thérapeutiques. La mission de la clinique n'est pas que le patient voie mieux, mais qu'il voie plus efficacement et à moindre effort, afin de renforcer considérablement son autonomie.

L'observation du fonctionnement des différentes unités de rééducation, cf. Figure 2.20, m'a permis de cerner les informations nécessaires pour comprendre les besoins des malvoyants ayant une vision très faible et de comprendre les techniques qu'emploient les différents spécialistes de l'institut. L'objectif sera par la suite de prendre en compte ces techniques dans nos travaux de développement d'un système d'assistance pour les malvoyants. La clinique propose plusieurs programmes adaptés aux besoins des malvoyants ou des non-voyants. Ces programmes sont déclinés, adaptés en fonction des besoins de chaque patient. La clinique comprend différentes unités de rééducation qui se complètent pour tenter de redonner, ou réapprendre, au malvoyant la perception de l'espace, la perception du corps dans cet espace, ainsi que de fournir des outils pour l'exploration de l'espace tel que le maniement de la canne blanche.

Je vais présenter succinctement et dans les grandes lignes, les éléments de rééducation de chaque unité.

³⁷ L'institut ARAMAV est un centre, basé à Nîmes, de réadaptation pour personnes malvoyantes ou non-voyantes. Le sigle ARAMAV signifie association de réinsertion des aveugles et malvoyants. L'AGE ARAMAV est l'association de gestion de l'établissement Institut ARAMAV, également dénommé Clinique du belvédère.

- La première est l'unité d'orthoptie, chargée d'optimiser l'utilisation de la vision résiduelle.
- La seconde est l'unité d'ergothérapie, chargée d'apprendre et d'optimiser les actions et les déplacements des patients lors des tâches quotidiennes telles que la prise de repas, le bricolage, la lecture ...
- La troisième est celle de la rééducation en psychomotricité, qui a pour objectif l'exploitation du travail en orthoptie et des sens compensatoires pour développer l'autonomie des patients dans les déplacements.



Figure 2.20 : Organisation de l'ARAMAV

Les différentes unités visent à rendre la prise d'informations visuelles par les patients plus efficace. Ces informations collectées se répartissent en deux catégories :

- les repères, en faible nombre, permettant la construction de la carte mentale du patient : représentation de l'espace environnant basé sur des signes visuels. Ils se caractérisent par des éléments discernables, caractéristiques et stables du lieu, tels que la hauteur, la largeur de la pièce, de la zone.
- les indices : caractéristiques secondaires pour se repérer et comprendre l'environnement, pas forcément d'ordre visuel (sens de circulation, sens de stationnement des voitures, ligne de végétation, bord de trottoir...).

Le temps d'analyse de la scène doit être réduit au minimum tout en permettant une prise d'informations pertinentes. Les différentes unités de rééducation soulignent l'importance d'une motricité oculaire combinée à un balayage de la tête. Cette stratégie implique une prise

d'informations rapide, ainsi qu'une exploration active du champ visuel, selon une stratégie précise.

2.3.1.2. L'orthoptie

Cette unité de rééducation s'adresse à des patients ayant encore un potentiel visuel. Elle n'améliore pas l'acuité visuelle, n'aggrave, ni ne stabilise la maladie. Elle consiste en une stimulation visuelle et permet de mettre en place de nouvelles stratégies de compensation du déficit visuel. Dans un premier temps, cette unité réalise un bilan précis des caractéristiques de la vision résiduelle. Puis, l'orthoptiste réalise la rééducation à proprement parler. Ses objectifs sont :

- L'exploration visuelle de l'espace : il s'agit d'étendre le domaine d'exploration oculaire en utilisant les techniques de balayage, fixation et saccade. Les personnes malvoyantes ont parfois tendance à conserver un regard fixe, dans ce cas, elles sont pénalisées par le processus d'acquisition des informations. En effet, en réalisant peu de mouvements oculaires, ou des mouvements de faible amplitude, on a alors accès qu'à une petite part du champ visuel. Afin d'accéder à toute la zone du champ visuel, il est nécessaire de pouvoir balayer toute la zone par des mouvements oculaires. Il est nécessaire que le patient suive un entraînement afin de pouvoir réaliser ces mouvements. Lors des autres étapes de la rééducation, cette faculté est indispensable. Lorsque le patient conserve le regard fixe, il se prive d'une quantité très importante d'informations.
- La localisation d'un élément ou d'un groupe d'éléments reconnaissables : il s'agit d'utiliser au mieux la vision résiduelle du patient et de développer une stratégie de recherche sélective de l'information visuelle.
- L'« interaction fonctionnelle » par le patient : il s'agit de développer l'aptitude du patient à générer l'hypothèse la plus probable sur la nature de ce qui est observé et de récupérer les informations non-accessibles visuellement par le raisonnement. En d'autres termes, le patient doit utiliser son raisonnement pour guider son exploration visuelle en fonction des informations qu'il vient de recueillir et les interpréter. Par exemple, si le patient a identifié un visage, il est très fortement probable qu'il soit face à une personne, dans ce cas, il a besoin de faire des mouvements oculaires proches pour analyser le reste de la personne afin de mieux la reconnaître. De même, si le patient a reconnu des lettres, il est a priori face à un document à lire, il doit alors

adapter sa stratégie d'analyse de la scène en fonction des éléments identifiés et se servir de ces informations pour ne pas avoir à analyser toute la scène.

- La rééducation du patient : il s'agit de donner au patient les moyens d'être capable de repérer certains indices visuels spécifiques comme les sources lumineuses et de reconnaître certaines informations (plan, indications visuelles, lettres, images, etc.).

L'orthoptie a pour objectif d'évaluer les potentiels visuels restants, de rééduquer un patient en lui permettant d'utiliser de manière efficace ses capacités et donc de collecter un maximum d'informations à la fois. Pour cela, le patient doit apprendre à utiliser ses mouvements oculaires et à développer son endurance visuelle, afin d'être efficace lors de la collecte d'informations.

2.3.1.3. L'ergothérapie

Le but de cette unité est de développer les sens compensatoires de la vision en se basant sur le toucher, l'ouïe, l'odorat et le goût pour acquérir un maximum d'autonomie. L'ergothérapeute réalise tout d'abord une analyse des capacités actuelles du patient. Il analyse la capacité à collecter des informations de type taille, texture, couleurs, poids, forme. Puis il met le patient en situation, afin de réaliser des activités globales (e. g. travaux manuels) et des activités de la vie journalière telle que la cuisine, l'écriture, la reconnaissance de billets et pièces, etc. Le patient réalise des activités de communication telle que le braille, la dactylographie et l'informatique.

Pour cela, l'ergothérapeute commence par rééduquer, puis réadapter avant de réinsérer le patient dans son cadre de vie habituel. Il est important de garder en mémoire qu'il est possible de « rééduquer seulement quelque chose de fonctionnel », c'est-à-dire que la personne doit encore avoir une certaine vision. Si cela n'est plus le cas, il est directement effectué la phase de réadaptation.

Lors de la rééducation, l'ergothérapeute analyse la bonne prise en compte de prérequis, c'est-à-dire la collecte des informations en utilisant les techniques apprises lors des séances d'orthoptie. Il réalise ensuite une phase d'utilisation de la mémoire et des sens cognitifs lors d'exercices simples.

Le patient doit développer sa coordination « œil-main » avec et sans les aides techniques tels que les loupes, les télé-agrandisseurs. Pour cela, le patient doit être capable de garder le contrôle de l'aide technique, l'objectif étant que le patient puisse maîtriser cette coordination

en toute situation. Cette coordination permet à terme au patient de réaliser des gestes de manière directe, sans hésitation et avec précision.

Le patient apprend à observer une scène environnante en utilisant son regard « visuo - tactile », c'est-à-dire qu'il utilise le toucher pour comprendre la scène et identifier les objets.

Une fois que ces présrequis sont acquis, une phase de réadaptation permet de mettre en situation le patient qui doit utiliser tous les prérequis qu'il a appris afin de réaliser des tâches de la vie quotidienne telle que la cuisine, du bricolage, des jeux ...

La phase de réinsertion, correspond à la phase finale, l'ergothérapeute va alors analyser l'environnement dans lequel le patient évolue normalement (son domicile, son travail, ...) afin de pouvoir l'adapter si besoin. Ce travail vise à permettre au patient de vivre avec sa déficience visuelle et de réaliser les activités de la vie quotidienne avec ou sans les aides techniques. Le patient doit pouvoir retrouver une vie proche de celle qu'il avait avant l'apparition de sa déficience en utilisant les méthodes qu'il a apprises et en utilisant les dispositifs d'assistance qui sont à sa disposition.

2.3.1.4. La psychomotricité et la locomotion

Cette unité a pour rôle d'améliorer la psychomotricité et la locomotion du patient :

- La psychomotricité est basé sur le travail corporel pour pallier les difficultés engendrées par la déficience visuelle, telles que les troubles de l'équilibre et de la perception de l'espace. Il s'agit de mieux appréhender son corps dans l'espace (notion de haut, bas, gauche et droite) et de développer une représentation mentale de l'espace: les objets ne sont plus simplement perçus, mais également localisés par rapport au sujet, le patient doit avoir conscience des relations spatiales entre les objets. Un travail important est effectué sur l'équilibre : ce qui inclut un travail sur l'utilisation de l'horizon visuel, de la compensation de l'information visuelle par l'information non-visuelle et de l'alignement corps / tête.
- L'objectif de la locomotion est de permettre à la personne de se déplacer et de percevoir le monde environnant dans une bonne posture de manière sûre et autonome avec un maximum de confort et d'aisance.

Un travail important pour la mobilité est l'estimation de la distance dans deux situations :

- Immobile : le patient qui est dans une position statique (sans se déplacer), doit déterminer sa position par rapport à des objets, des obstacles mobiles ; il doit détecter et expliquer comment éviter les obstacles lors de son déplacement.
- en déplacement : le patient doit se rendre à une destination précise sans entrer en collision avec des objets, pour cela, il doit les identifier et les localiser.

Lors de la mobilité, les patients ont besoin d'identifier leur trajet, identifier la taille de la zone de déambulation et sa délimitation. Il faut également pouvoir identifier les obstacles et les dangers. Il est nécessaire de gérer l'espace, d'anticiper pour éviter les obstacles, d'identifier les passages piétons, le bord des trottoirs, les marches... Les patients commencent à analyser l'endroit où ils se trouvent afin de prendre des repères, d'identifier les bâtiments, les sens de circulation et de stationnement. Il est nécessaire d'identifier les quilles, les panneaux publicitaires, les panneaux de signalisation, les voitures, le stationnement et la circulation. Pendant les phases de déplacement, les patients doivent pouvoir analyser la scène afin de déterminer la largeur et la position de la voie de circulation qui leur est accessible. Une partie de ce travail est basée sur la négociation de franchissement de type marches, escaliers (la première et la dernière marche suffisent), l'utilisation d'indices indirects : mouvement des piétons, localisation d'une rampe adaptée.

Lors de phase de mobilité, les malvoyants et non-voyants utilisent souvent une canne blanche ou un chien-guide. Le rôle de la canne est tout d'abord de dégager le regard du sol. Lorsque l'on passe d'une surface à une autre, la canne permet l'estimation de la pente d'arrivée par rapport à la pente précédente, ainsi que le dénivelé et l'état de surface (lisse, rugueux, avec des trous...), cf. Figure 2.21 à droite [14]. La canne permet d'analyser l'espace situé devant le malvoyant. Il est possible de détecter un passage piéton par la détection de quilles, ou d'une bande podotactile, comme sur la Figure 2.21 à gauche. L'analyse des différentes textures présentes, comme sur la Figure 2.21 à droite et sur la Figure 2.22, permettent de détecter les zones de la rue prévues pour les piétons, les voitures, les transports publics... Il est également possible de détecter les changements de niveau dans une rue, un couloir, cela permet à la personne d'identifier le bord d'un trottoir, cf. Figure 2.22, ou d'obtenir la largeur d'un trottoir lorsque celui-ci est étroit. Toutes ces informations normalement accessibles à une personne bien voyante sont essentielles pour un malvoyant afin qu'il puisse utiliser, emprunter un passage, ou un trottoir de manière sûre, sans se mettre en danger. De même, l'analyse des dénivelés permet de détecter les escaliers montants ou descendants [14].



Figure 2.21 : Exemple de passage piéton avec repère podotactile (à gauche) et une place comportant différentes textures de sol (à droite)³⁸.



Figure 2.22 : Exemple de trottoir large (à gauche) et de trottoir étroit (à droite)³⁹.

Comme le montre la Figure 2.23, la zone de l'espace analysable avec une canne blanche est très limitée. La canne blanche ne permet pas de détecter des obstacles situés en hauteur tels que des panneaux de signalisation.

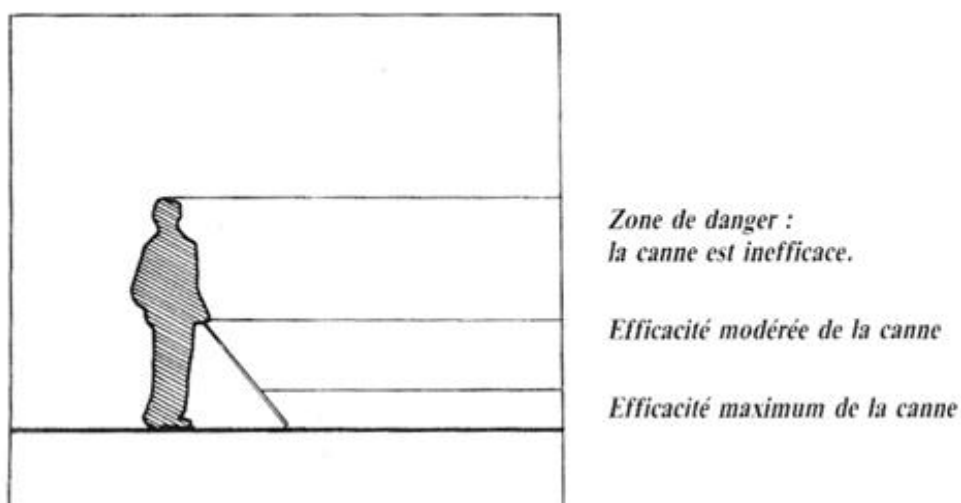


Figure 2.23 : Zone explorable avec une canne blanche⁴⁰.

³⁸ <http://tel.archives-ouvertes.fr/tel-00593859> [14].

³⁹ <http://tel.archives-ouvertes.fr/tel-00593859> [14].

Le malvoyant peut adapter sa posture en prévision de la zone sur laquelle il doit se déplacer. Enfin, la canne blanche est aussi un symbole indiquant aux autres personnes que son utilisateur présente une déficience visuelle.

2.3.2. Analyse

J'ai présenté dans cette partie, une synthèse des discussions avec les professionnels de la réadaptation de cette clinique et les patients. Comme nous venons de le voir, les capacités visuelles des malvoyants ne sont pas sollicitées dans un contexte de mobilité. Il peut techniquement être réalisé une assistance visuelle des malvoyants, comme cela est présenté dans la section 1 du chapitre 3. Je vais donc présenter une synthèse des besoins des malvoyants et les recommandations des professionnels pour le développement d'une aide visuelle.

Tout d'abord du point de vue physiologique, il est important d'après les professionnels interrogés, de laisser le patient chercher l'information, car il faut garder les mouvements oculaires libres. En effet, si l'affichage n'est pas réalisé dans de bonnes conditions, le patient pourrait alors limiter ses mouvements oculaires afin que les informations se posent dans son champ visuel ; et à terme ne plus faire de mouvements oculaires. Il va ainsi perdre une partie de sa capacité à réaliser les mouvements oculaires qui lui seront pourtant utiles lorsque le système d'aide ne pourra pas lui apporter un soutien.

De même, certains des professionnels interviewés ont indiqué qu'il faudrait afficher les informations leur permettant d'extraire plus facilement les indices visuels utilisés, à savoir, les contours, les contrastes et les couleurs :

- pour les malvoyants avec une atteinte de la vision périphérique, il serait intéressant de réduire l'image afin de mettre le plus d'informations possibles dans le champ fonctionnel sans le saturer et en tenant compte de l'acuité résiduelle de la personne,
- pour les malvoyants atteints de dégénérescence du centre de la rétine, il faudrait réaliser un grossissement et/ou une modification des contrastes afin que la zone périphérique de la rétine puisse néanmoins transmettre des informations utiles.

Certains professionnels ont également indiqué qu'il est important de laisser le malvoyant maîtriser l'aide technique et de ne pas faire une aide technique qui remplace le malvoyant. Le

⁴⁰ <http://tel.archives-ouvertes.fr/tel-00593859> [14].

patient serait perturbé lorsque celle-ci ne pourra pas lui être utile. De plus, cela lui impose de rester actif et de garder le contrôle de la situation.

Pour les professionnels de la psychomotricité et les instructeurs en locomotion, il est important qu'une aide visuelle permette de conserver l'équilibre et ne perturbe pas les autres sens. Il est également important que ce type d'aide se répartisse sur la personne sans créer de gêne, de déséquilibre. Le poids doit être le plus faible possible et être réparti de manière équilibrée afin de ne pas modifier les sensations que le malvoyant a appris à percevoir avec son corps. Il est donc primordial d'avoir un bon positionnement des informations affichées par rapport à la scène, et d'utiliser le son avec modération, car le patient sollicite les indices sonores pour conserver son équilibre.

Les malvoyants ont exprimé, tout comme les professionnels, des besoins : de simplicité d'utilisation des aides existantes, telles que les loupes et télé-agrandisseurs ; d'obtenir des images agrandies plus simples ; et surtout d'améliorer les outils actuels afin de leur fournir des « contrastes mieux détectés ». Autrement dit, les patients souhaitent des images « nettoyées » et des contrastes affichés correspondants aux contours des pièces ou des éléments importants de la scène.

Il est souvent nécessaire pour les malvoyants d'avoir une loupe ou une loupe électronique, un télé-agrandisseur à proximité. Au regard du poids et de la taille de ces appareils, il n'est pas facile de les réunir tous ensemble à proximité en permanence. Certains professionnels suggèrent d'intégrer une loupe dans un dispositif portable dans le but de réduire le nombre d'équipements à utiliser ou à transporter par le patient.

Lors de discussions avec les malvoyants, ceux-ci ont exprimé des besoins afin de :

- Regarder à l'extérieur.
- Reconnaître des personnes : qu'ils rencontrent, croisent, afin de les identifier. En fonction des malvoyants, certains souhaitent soit une solution automatique de reconnaissance de visage, soit pour un rehaussement des visages permettant au patient de réaliser la reconnaissance lui-même.
- Identifier les couleurs : d'un objet, comparer deux couleurs pour réaliser par exemple des tâches de la vie courante, telles que choisir des vêtements assortis.
- Augmenter les contrastes : afin de mieux différencier des objets, de bricoler, marcher, différencier le bord des marches, différencier le bord des lettres sans les ombres.

- Augmenter ou diminuer la luminosité des images afin de compenser la luminosité de l'éclairage ambiant.
- Interagir et utiliser les appareils grand-public lorsque cela est possible, tels que Ipad, ordinateur, smartphone, via un clavier adapté (clavier azerty grossi ou clavier braille) et une mise en forme de l'écran (tel que les couleurs, le grossissement), ce qui leur permet de bénéficier de toutes les applications source de contenus, d'informations disponibles.
- Regarder la télévision : avec un grossissement ou avec modification des contrastes et des couleurs, sans perturber l'image pour toutes les personnes qui regardent la télévision avec eux. Il est également souhaité que cette fonctionnalité n'occasionne pas de problème de décalage du son par rapport au son du téléviseur.
- Interagir vocalement avec le dispositif sans saturer le sens auditif qui est utilisé pour maintenir l'équilibre et analyser l'environnement.

Certains malvoyants ont expliqué ne pas utiliser d'aide dans certaines situations de peur du regard des autres personnes envers leur problème de vue. Je peux notamment citer le cas des jeux de cartes avec un groupe d'amis où certains n'osent pas sortir une loupe ou un éclairage adapté pour jouer. Cela a déjà été souligné dans l'étude de Shinohara et Wobbrock [173], où il a été mis en évidence que certains patients n'utilisent pas les aides à disposition du fait du regard des autres personnes sur leur handicap.

Toutes ces informations sont très importantes pour réaliser, concevoir un dispositif d'assistance visuelle qui peut être accepté par les malvoyants dans la mesure où il répond à leurs attentes. En effet, la prise en compte de leurs besoins que ce soit en terme d'ergonomie, de poids, d'usage, de simplicité d'usage et de respect des besoins physiologiques, est très importante pour le succès d'une aide visuelle. Ces informations ont donc été prises en compte lors des développements relatifs au chapitre 3 concernant les développements algorithmiques et au sein du chapitre 5 relatif au développement d'un démonstrateur mobile.

3. Image et contours pour malvoyants

Dans ce chapitre, je vais commencer par faire un état de l'art des techniques permettant d'analyser les mécanismes de la vision, les modèles de saillances visuelles, puis les techniques de rehaussement d'image et de contours avant de présenter quelques dispositifs développés pour aider les malvoyants à mieux percevoir l'environnement. Cette présentation, dans la section 3.1, permettra de justifier mon choix de travailler sur le rehaussement de contours. Dans la section 3.2, je vais présenter d'une part des banques d'images existantes et celle que j'ai utilisée pour mes développements. Je vais ensuite présenter un certain nombre de méthodes de détection de contours et d'analyse multi-échelle de l'information, section 3.2.5. Dans la section 3.2.6, je vais présenter mon détecteur de contours en détaillant chaque élément et en justifiant chaque élément de celui-ci. Je vais également réaliser une comparaison des résultats de ce détecteur avec les autres détecteurs de contours présentés dans la section 3.2. Enfin, dans la section 3.3, je vais conclure cette partie relative à mes développements en traitement d'images.

3.1. Etat de l'art

3.1.1. Introduction

Plusieurs systèmes récents d'aides pour malvoyants, tels que le télé-agrandisseur, la loupe électronique, la canne blanche électronique, le GPS adapté, les systèmes vocaux et le gant tactile permettent de répondre à certains de leurs besoins en fonction de leurs capacités visuelles résiduelles (visuelle, oculomotrice, cognitive) [153]. Ces différentes techniques permettent de réaliser certaines tâches, cependant, il reste un grand nombre de besoins non satisfait à ce jour. Pour ces besoins, les avancées techniques récentes peuvent être mises à profit pour y répondre. Nous allons présenter dans cette partie un état de l'art des techniques de suivi du regard (oculométrie) et d'analyse des mécanismes de la vision, des outils de traitement d'images pouvant répondre à ces besoins, ainsi que des solutions permettant d'utiliser ces outils dans des systèmes dédiés aux malvoyants.

3.1.2. Oculométrie

Le système visuel humain réalise la collecte et l'interprétation des informations visuelles environnantes [198]. Pour cela, nos yeux réalisent différents mouvements oculaires tels que les saccades, poursuites, dans le but de couvrir une grande partie de notre champ visuel et suivre les zones de forte information. L'oculométrie est le nom donné à la technique qui consiste à collecter, enregistrer les positions des yeux dans le temps, puis localiser dans l'espace les lignes de regard, le point de fixation ou la zone observée par la personne [122]. Il est ainsi possible de collecter les positions des fixations, cf. Figure 3.1, d'analyser les saccades en tailles, en fréquences, en vision de près de manière statique ou en mouvement [52]. L'oculométrie peut être utilisée pour analyser les mécanismes de la vision.

Une fois les informations relatives aux points de fixations collectées on peut déterminer s'il s'agit d'une saccade, d'une fixation [174].

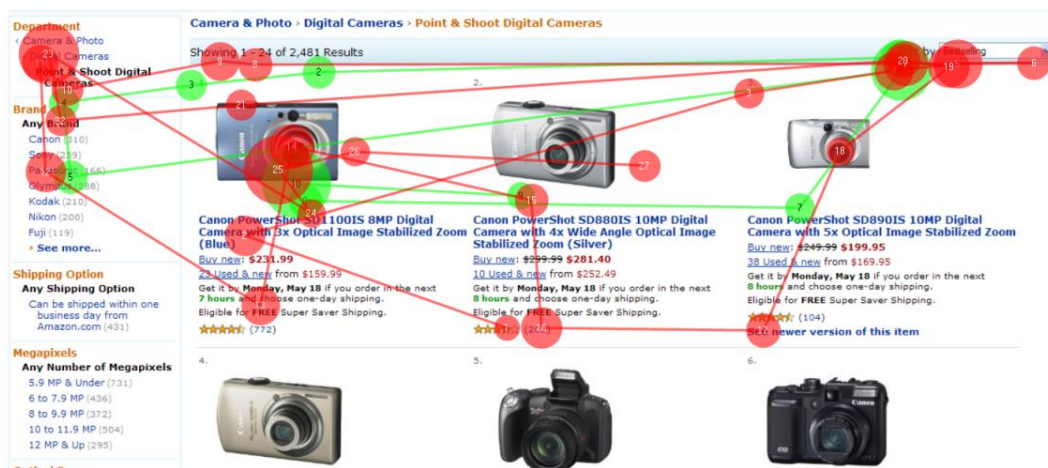


Figure 3.1 : Exemple de suivi de position du regard⁴¹

La réalisation d'un simulateur permet d'analyser les mécanismes de la vision de pouvoir comparer plus facilement la vision d'une personne bien voyante avec ou sans la présence d'un scotome. Il est important de pouvoir simuler avec suffisamment de performance et d'être en temps réel pour ne pas biaiser les mesures et analyses réalisées [6]. L'utilisation d'ordinateurs performants et des techniques de programmation de visualisation 3D permet de simuler la vision, avec ou sans scotome, de prendre en compte le flou périphérique et/ou la baisse d'acuité de la vision périphérique et de collecter la position des yeux par rapport à cette

⁴¹ <http://www.tobii.com/>

simulation [53, 51]. L'oculométrie nécessite de pouvoir traiter rapidement les données provenant du dispositif de suivi pour calculer la position de chaque œil.

Ces tests sur simulateur peuvent permettre la mise en évidence des mécanismes de la vision, de valider des choix matériels comme le type d'écran, la résolution choisie, le type d'affichage, le choix du mode d'affichage (noir et blanc, couleur, 3D ...) [131]. Cette technique a notamment permis de montrer qu'au-delà d'une certaine vitesse le système visuel n'est plus capable de collecter ces informations [66, 65], un même point observé à différentes fréquences ne sera pas perçu de la même manière. De même, il a été mis en évidence que les mouvements oculaires sont adaptés en fonction des besoins du système visuel [7].

L'analyse des mouvements oculaires permet de définir des modèles pour prédire les positions de fixation en fonction du gain d'information potentiel lors de chaque mouvement et de chaque position de fixation [67], ou pour développer des méthodes pour la rééducation des mouvements oculaires en affichant des stimuli dans les zones adéquates et de pouvoir contrôler le fait que le patient effectue bien ces mouvements. Il est également possible d'analyser le champ visuel, les mouvements oculaires [121]. Actuellement, il n'existe pas de dispositif suffisamment compact et performant pour réaliser cette l'analyse, excepté sur des simulateurs^{42 43}, sur une paire de lunettes, il est uniquement possible de réaliser la détection de la position du regard pour des interactions logicielles^{44 45 46}, cf. Figure 3.2.

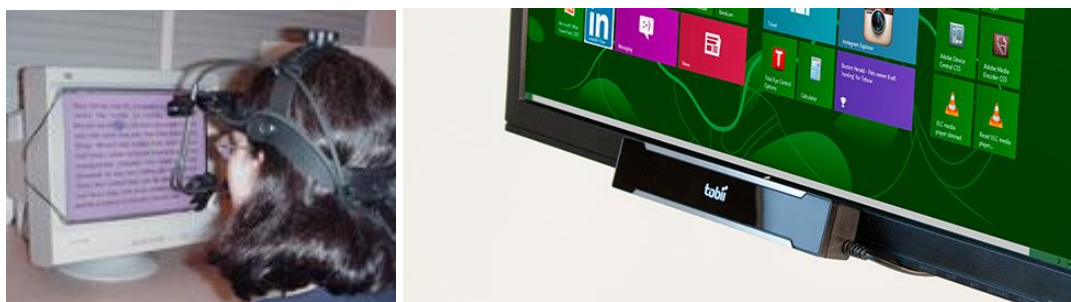


Figure 3.2 : Exemple de dispositifs d'oculométrie : (à gauche) un dispositif pour la recherche ; (à droite) un dispositif pour tablette et PC.

⁴² <http://www.sr-research.com/eyelinkII.html>

⁴³ <http://www.arringtonresearch.com/headfixed.html>

⁴⁴ <http://www.tobii.com/en/assistive-technology/global/products/hardware/tobii-eyemobile/>

⁴⁵ <http://www.tobii.com/en/eye-tracking-research/global/products/>

⁴⁶ <http://www.mygaze.com/>

3.1.3. Modélisation des mécanismes de la vision et assistance visuelle

L'analyse des mécanismes de la vision permet de comprendre le fonctionnement de ces mécanismes ainsi que les facteurs pouvant influencer leurs performances [39]. Ces analyses permettent notamment de mieux comprendre les mécanismes de la pseudo fovéa dit PRL (pseudo rétinale location), des mécanismes de recherche de l'information visuelle, pour la vision de près, la vision de loin, la détection des obstacles. Une fois ces mécanismes connus, il est possible de développer des dispositifs d'aide visuelle adaptés.

Comme nous l'avons vu au chapitre 1, certaines pathologies rétiniennes affectent le champ visuel du patient. Le champ visuel peut soit être rétréci : champ tubulaire, soit comporter un scotome : zone non ou peu fonctionnelle sur la rétine.

Avec un scotome dans le champ visuel, le patient n'a généralement plus l'usage de la fovéa. Il doit alors trouver un moyen de collecter le plus d'informations importantes possible. Pour cela, il développe généralement une PRL ou une pseudo-fovéa de fixation zone préférée avec laquelle le patient va fixer les objets afin d'être moins perturbé par son scotome [18].

Pour la lecture, le patient doit identifier des lettres dans un document [171, 172]. Il a été analysé un certain nombre de paramètres tel que le lien entre l'aisance de lecture et la taille de l'interligne [20], l'espace entre les mots et l'espace entre les lettres afin de faciliter la lecture [36, 21]. De même, il a été mis en évidence la relation entre la taille de la zone de fixation, la taille du scotome, la vitesse de lecture et la taille des saccades réalisées pour lire, cf. Figure 3.3 [168, 169, 19].

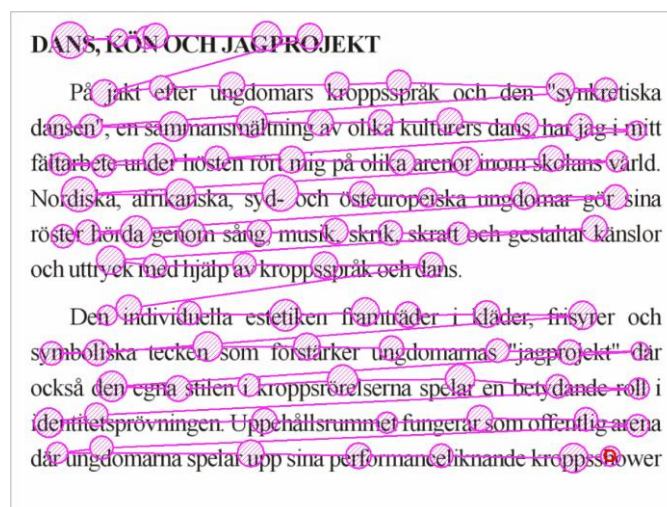


Figure 3.3 : Exemple de fixations et saccades lors de la lecture.

Une fois la compréhension des mécanismes de la vision réalisée, il est possible de mettre en forme les informations dans le but d'aider les malvoyants à lire des documents numériques [22, 114] ou papier. L'identification des zones importantes d'un document permet d'aider le patient, en mettant en évidence ces éléments, pour parcourir le document [166] [167]. Il peut être défini la notion de ligne de regard, ligne indiquant la zone analysée par l'œil. Elle est définie par le centre de la fovéa ou la PRL et le centre de l'œil. La connaissance de la position des lignes de regard, en temps réel, permet de placer les informations très utiles dans des zones fonctionnelles du champ visuel de la personne [170].

Pour l'analyse de ces mécanismes, il est souvent utilisé des simulateurs de lecture avec un scotome, de marche, de conduite. Il est possible de mettre en évidence les besoins des personnes malvoyantes et de comprendre les mécanismes utilisés par le système visuel. Le patient cherche à détecter les passants à éviter, les autres voitures dans son sens et venant dans le sens opposé, les feux et la signalisation [145, 74]. Au lieu de réaliser une simulation sur écran et un patient assis devant celui-ci, il peut être nécessaire de développer un simulateur de marche ou de conduite. La personne va solliciter des mécanismes proches de ceux utilisés pour la lecture de manière à extraire les données essentielles à sa perception et son repérage dans l'espace [150, 186].

La qualité de la simulation est importante pour une modélisation fidèle des mécanismes. La simulation permet de ne pas exposer inutilement des malvoyants à des risques, par exemple pour l'analyse de l'aisance en mobilité, cela permet de ne pas faire traverser une rue bondée ou mal éclairée à la personne. Cependant certains simulateurs peuvent influencer les analyses et les résultats obtenus, s'ils ne sont pas suffisamment fidèles à la réalité [196, 15] ; dans de tels cas, les modèles des mécanismes de la vision peuvent être erronés ou incorrects.

3.1.4. Modèles de saillance et d'apparence visuelle

Au début des années 80, David Marr, avait proposé une théorie de la vision introduisant les grands principes des neurosciences et des modèles d'apparences visuelles [112]. Cette théorie propose toute une série de principes, en se basant sur les mécanismes de la vision, devant permettre la compréhension et l'interprétation des images.

Treisman et Gelade ont développé le « feature theory of attention » [183] en même temps que Marr. Ce modèle commence par la décomposition d'une image en plusieurs cartes de

paramètres tels que la couleur, la luminance, les mouvements afin de pouvoir les visualiser et les analyser séparément.

Ce modèle inclut une modélisation de deux mécanismes d'analyse utilisée par le système visuel, le « bottom-up » réalisant une analyse bas niveau de l'image en collectant les zones saillantes, à fort contenu informationnel et le « top-down » réalisant une analyse haut niveau de la scène en collectant les informations additionnelles, non détectées par le « bottom-up » pour certaines tâches. Cette deuxième branche est un élément important du système visuel, car elle lui permet en fonction de la situation de compléter les informations nécessaires à la compréhension de la scène. Par exemple pour la reconnaissance de visage, le mécanisme « bottom-up » va permettre de collecter les points essentiels du visage nous permettant d'identifier le visage sans le reconnaître, le mécanisme « top-down » permettra de collecter les informations nécessaires pour le reconnaître.

Wolfe avait également proposé un modèle ressemblant fortement au fonctionnement de la rétine en modélisant les cônes, les bâtonnets, les cellules horizontales et bipolaires et les mécanismes « bottom-up » et « top-down » [189].

Sur le même principe que les modèles précédents, Laurent Itti a proposé un modèle considéré comme une référence parmi les modèles d'apparence visuelle [88, 85, 86, 84]. Ce modèle donne plus d'importance aux pixels contenant de l'information et tente de réduire le poids des pixels présentant peu d'information [87, 152]. Ce modèle intègre également un mécanisme de prédiction des erreurs de fixation afin de minimiser les erreurs possibles [190]. La carte de saillance est générée sur le principe du « winner take all », c'est-à-dire que la carte d'information de plus forte valeur remporte sur toutes les cartes [132, 32].

L'utilisation de la couleur permet également d'améliorer ces modèles. La prédiction des zones d'intérêt dans une image, cf. Figure 3.4, permet de réaliser des segmentations d'images ou de la reconnaissance de visages [163, 98, 138].

Ces modèles produisent des cartes de saliences indiquant les zones d'intérêt des objets à conserver lors de segmentation [5, 4, 3, 43]. La détection de l'objet comportant le plus de contenu informationnel peut être intéressante pour l'assistance aux non-voyants et malvoyants, on peut ainsi identifier l'objet potentiellement observé et leur apporter des informations relatives à cet objet [180]. Cependant, ces modèles ne permettent pas toujours de détecter des objets complexes comprenant plusieurs éléments de teintes différentes.



Figure 3.4 : (à gauche) Image d'origine ; (à droite) reconstruction de l'image en conservant uniquement les zones saillantes de l'image [48, 47]

Ces modèles permettent d'obtenir de l'information souvent localisée dans quelques zones de l'image, il serait intéressant pour l'assistance aux malvoyants d'obtenir de l'information dans plusieurs zones de l'image. La connaissance des zones d'intérêts de l'image, de la scène peut permettre de déformer l'image ou de mettre en forme les informations pour les transmettre à la personne dans les zones de son champ visuel utile.

3.1.5. Rehaussement et anamorphose d'image

La connaissance de la forme du scotome ou des déformations du champ visuel permet de réaliser une déformation de l'image afin de mettre les informations en dehors des zones non-fonctionnelles de la rétine [58]. Il existe un grand nombre de déformations possibles de l'image telle que les anamorphoses : déformation réversible d'une image afin de la mettre à l'échelle, de corriger une perspective [164], ou les déformations permettant de déplacer les informations situées derrière le scotome de la personne.

L'anamorphose de la scène peut également être réalisée avec des prismes colorés pour les personnes atteintes de rétinite pigmentaire [192, 194, 143, 70] ou des prismes non teintés est également possible [77]. Cette solution optique permet de transmettre des informations à des patients atteints de pertes du champ visuel [162], cependant, cela nécessite que le patient conserve une bonne acuité pour pouvoir utiliser ce type de dispositif. Les solutions d'anamorphoses des images par système optique peuvent répondre à différents usages en fonction des législations des états. Ces systèmes optiques pour malvoyants peuvent être

utilisés pour conduire un véhicule plus aisément en fonction de la législation des pays où vivent les patients, par exemple certains états des Etats-Unis d'Amérique [29, 136].

De même le codage en DCT (Discrete cosine transform), ou transformation en cosinus discrète, utilisé dans les formats JPEG et MPEG⁴⁷ permet de réaliser un rehaussement, en modifiant les coefficients des matrices de divisions ou multiplications [199], pour la lecture, l'observation de scène et les vidéos [151, 57]. La DCT est la décomposition d'un sous-bloc de l'image en une matrice, cf. Figure 3.5, généralement de taille 8 x 8, codant la valeur correspondant à un cosinus précis pour chaque élément de la matrice ; cela permet de réaliser une décomposition des basses fréquences de l'image. Cet outil est très utilisé dans les algorithmes de compression d'images et de flux vidéo, car il permet de conserver uniquement les basses fréquences de l'image.

La zone bleue sur la Figure 3.5 représente les fréquences spatiales à rehausser pour aider les personnes malvoyantes à mieux percevoir les informations. Ces fréquences sont les fréquences spatiales pour lesquelles le système visuel humain est le plus sensible.

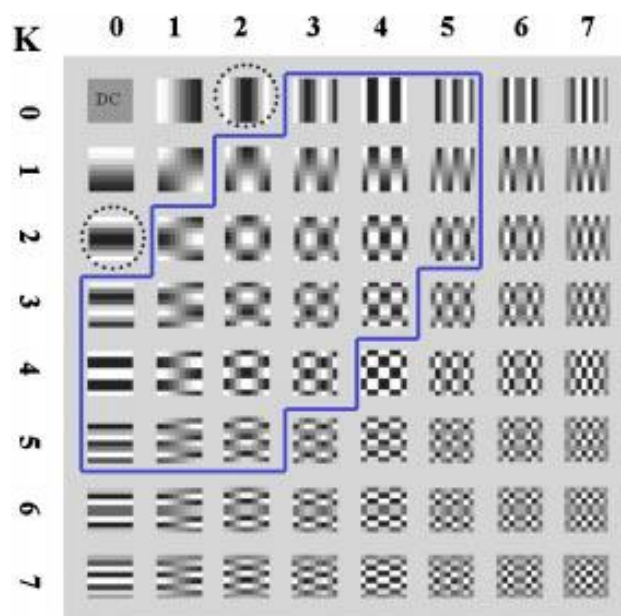


Figure 3.5 : Représentation des différentes fréquences d'une matrice DCT

Un dispositif permettant de rehausser un flux vidéo pour une personne malvoyante a été développé par l'équipe de Peli [8, 11, 9], cela permet de mettre en évidence le bord des lettres ou rehausser certaines fréquences dans le but d'aider les malvoyants [63], cf. Figure 3.6. Ce

⁴⁷ <http://www.bretl.com/mpeghtml/mpeg2vc1.htm>

rehaussement d'image présente l'inconvénient de s'appliquer pour toutes les personnes regardant la télévision [93, 135].

En combinant l'anamorphose et le rehaussement, il a été développé des systèmes semblables aux télé-agrandisseurs proposant des facteurs grossissants ajustables et permettant d'afficher en couleurs réelles, en fausses couleurs, en niveau de gris, en fil de fer [73].



Figure 3.6 : Rehaussement MPEG : (à gauche) l'image rehaussée ; (à droite) l'image d'origine en symétrie

De même, en analysant les fréquences spatiales, il est possible de détecter ou rehausser les contours des images. Dans ce cas, le rehaussement n'est pas appliqué à chaque élément de la bande de fréquences mais uniquement lorsque cet élément est considéré en plus comme un contour.

3.1.6. Détection de contours

John Canny a proposé un algorithme capable de détecter dans certaines conditions les contours de manière efficace [38, 37]. Ces dernières années, il a été proposé un certain nombre d'optimisations [27], et d'implémentations par ondelettes [83], méthodes réalisant une décomposition entre les basses et hautes fréquences de l'image pour accélérer le traitement. De même certains détecteurs de contours ont été optimisés et implémentés sur un système électronique tel qu'un FPGA [119, 110].

L'équipe d'Eli Peli a démontré que les contours combinés à un HMD permettent d'apporter une aide non-négligeable aux personnes malvoyantes [146, 149, 187]. Pour cela, ils ont développé leur propre détecteur de contours le « wide-band enhancement » [142, 144], mais

celui-ci n'est pas adapté à la réalité augmentée sur verres transparents [9]. En effet, cet outil rehausse les contours en utilisant aussi bien des éléments noirs ou blancs, or les systèmes de réalité augmentée à verres transparents, cf. Chapitre 5, ne peuvent pas restituer des couleurs sombres (celles-ci sont transparentes).

3.1.7. Dispositifs d'affichage sur tête

Un système d'affichage sur tête dit HMD (Head Mounted Display) [106, 187] peut permettre d'aider des personnes malvoyantes. Ces dispositifs sont présentés plus en détails au chapitre 5. Ce type de dispositif, de réalité augmentée à verre transparent permettant d'afficher des informations incrustées sur la réalité, est utilisé pour développer de futurs dispositifs d'assistance visuelle dédiés aux malvoyants. Le multiplexage (superposition de signaux, ou d'informations) d'un signal vidéo avec la réalité perçue, en affichant sur la réalité des informations virtuelles, sur un HMD permet d'ajouter des contours, cf. Figure 3.7, une image agrandie ou miniaturisée, un ajout localisé d'information [147, 140, 149].

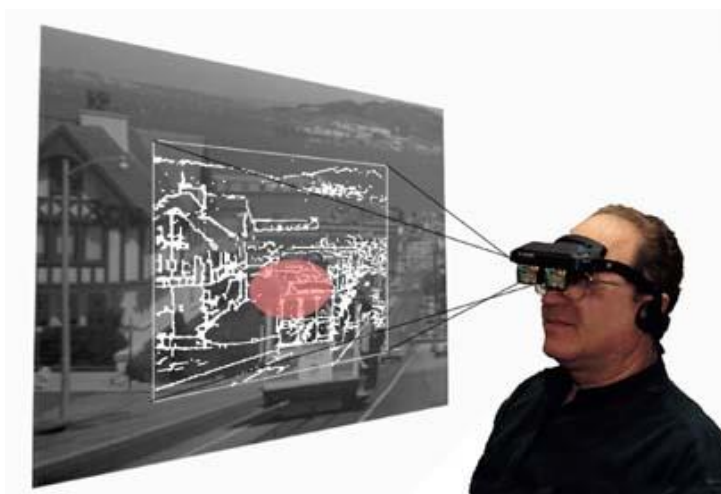


Figure 3.7 : Exemple de HMD, pour la restitution de contours, développé par l'équipe d'Eli Peli.

Il est ainsi possible de rehausser lorsque l'on a identifié les informations les zones d'intérêts [107] ou les informations très utiles [79], de les mettre en évidence, de les rehausser pour permettre aux patients de les analyser plus rapidement. Il est donc important de définir et de pouvoir détecter les zones d'intérêt dans une image. Cette analyse permet d'enrichir uniquement les informations les plus utiles pour les malvoyants [72, 75, 141, 175, 124, 123]. Une fois les informations détectées, il est alors possible de rehausser uniquement ces informations, afin de ne pas saturer le champ de vision du malvoyant d'information.

Lorsque l'on souhaite rehausser uniquement les informations, les plus pertinentes, pour le patient, il est intéressant de prendre en compte le type d'atteinte du champ visuel de celui-ci. Par exemple, un patient atteint d'une vision tunnel, c'est-à-dire une atteinte de la vision périphérique, doit réaliser beaucoup plus de mouvements oculaires afin de collecter les mêmes informations. Il est intéressant de faire plus ressortir les informations afin de lui permettre de les collecter plus aisément et plus rapidement [97] ou de les miniaturiser afin de mettre plus d'informations dans son champ visuel résiduel. L'équipe d'Eli Peli, de l'Université de Harvard, a notamment développé un modèle, d'apparence visuelle pour rehausser les images pour les malvoyants [144, 142].

De même en réalisant un rehaussement d'image sur un HMD suffisamment ouvert, permettant de bien percevoir l'environnement en dehors de la zone enrichie en informations, il est possible de développer un système de déplacement en vision nocturne [28, 30].

L'intégration de ces méthodes dans un HMD pour une personne malvoyante peut lui permettre également d'améliorer la détection des obstacles, l'évaluation de risque de collision et de naviguer librement sans risque [193, 148, 109]. L'usage de ces techniques dans un HMD peut dans certains cas engendrer des difficultés pour le patient lors de la collecte des informations affichées avec le système tête haute [10].

Un certain nombre d'études [146, 187] ont montré qu'il est important de permettre au malvoyant

- d'utiliser ses capacités visuelles résiduelles
- de conserver des mouvements oculaires naturels
- de fournir des informations sur le contenu de l'espace au centre et en périphérie de son champ visuel en fonction de sa pathologie
- de pouvoir utiliser la solution en toute luminosité de l'environnement ambiant
- de pouvoir utiliser sa correction ophtalmique
- de permettre un usage prolongé
- de régler, appréhender et adapter le système à la situation aisément
- d'avoir un usage confortable, avec un dispositif ergonomique lumineux et contrasté
- de conserver une mobilité aisée : système léger et de petite taille, design et ergonomique

Au vu des éléments présentés, il semble intéressant de pouvoir utiliser la détection de contours pour aider les personnes malvoyantes. Cette solution peut être réalisable et peut répondre aux besoins des personnes malvoyantes. De plus, cette solution a été testée sans une méthode de détection des contours adaptée, ce qui a tout de même permis de montrer son intérêt. Je vais maintenant présenter les outils utilisant des cartes de profondeur.

3.1.8. Orthèses visuelles et tactiles

Il est possible de d'analyser l'environnement avec des caméras et de transmettre les informations ainsi extraites aux patients qu'ils soient malvoyants ou non-voyants. Pour cela, il a été développé un dispositif visuo-tactile, qui comprend une paire de lunettes pour supporter la caméra, un système de calcul et un dispositif de restitution tactile, cf. Figure 3.8. L'image de l'environnement est acquise par une caméra stéréoscopique afin de disposer à la fois de l'information de profondeur et de couleur. Les informations sont restituées à la personne par le dispositif de restitution tactile, les picots de ce dispositif sont pilotés de manière à dessiner en relief sur celui-ci les informations à transmettre au patient [59, 92, 80].



Figure 3.8 : De gauche à droite: une image de l'environnement, la représentation tactile de l'environnement, et l'interface tactile (TactiPad, ISIR, Université Paris 6).

De même, il est possible de transmettre les informations par électronique ou optique comme cela est réalisé dans le cadre du projet LASDOP, cf. Figure 3.9, développé à Nîmes et à Montpellier [155, 179]. L'image est acquise par une caméra stéréoscopique afin de disposer à la fois de l'information de profondeur et de couleur. Cette représentation de l'environnement en couleurs et en profondeur permet de réaliser une simplification de la scène. Seules les informations simplifiées sont transmises au malvoyant.



Figure 3.9 : (à gauche) Une scène réelle ; (à droite) la restitution de cette scène par le stimulateur rétinien.

Dans ces deux systèmes, la carte de profondeur permet de fournir des informations utiles aux personnes malvoyantes. Cependant, ces informations sont restituées soit par un système tactile soit par un système de réalité virtuelle. Une carte de profondeur restituée sur un dispositif à verres transparents risque de remplir tout le champ de vision d'information et de masquer la réalité.

3.2. Détecteur de contours adaptatif

3.2.1. Banque d'images

3.2.1.1. Banque d'images « contours »

Nous avons besoin pour valider nos travaux de recherche d'une banque d'images contenant si possible une carte de contours fidèle pour les malvoyants.

L'université de Berkeley a créé une banque d'images⁴⁸ de test pour la détection d'objets et la segmentation [113]. Cette banque d'images existe en deux versions la BSDS300 et BSDS500. Ces deux banques respectivement de 300 et 500 images, contiennent les images originales en couleurs, en niveau de gris, les cartes contours pour au moins 5 observateurs et une carte

⁴⁸ <https://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>

fusionnant les réponses de ces observateurs. Cette banque d'images est très souvent utilisée pour comparer des outils de détection de contours, de segmentation d'images.

Lorsque l'on analyse les Figures 3.10 3.11 et 3.12, contenant chacune une image d'origine en couleur, deux cartes de contours réalisées par des observateurs bien voyants et une carte de fusion des cartes de contours des 5 observateurs ayant répondu pour chaque image de la banque. Sur l'image 3.10 les deux observateurs donnent des réponses très différentes pour décrire les bâtiments.

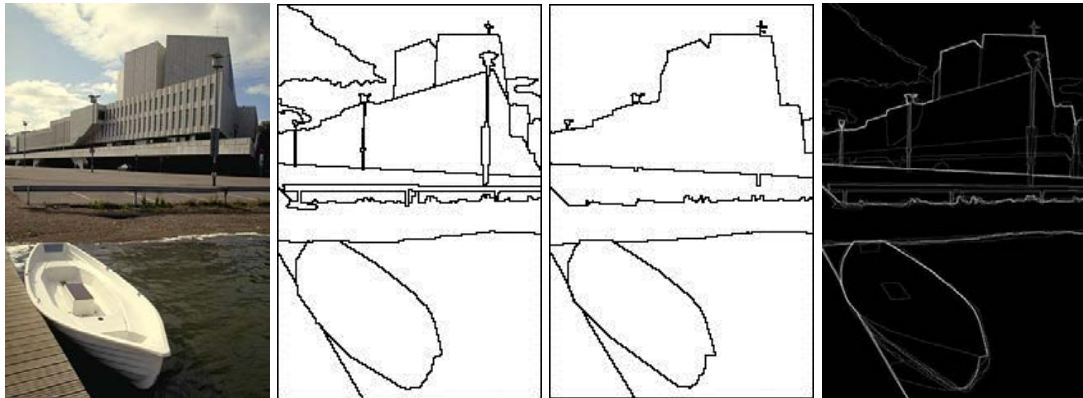


Figure 3.10 : Images de BSDS300, de gauche à droite : l'image originale en couleur, la carte de contours d'un observateur, la carte de contours d'un deuxième observateur, et la fusion des cartes de contours des observateurs

Sur l'image 3.11, un des observateurs uniquement indique le chemin et les maisons complètement. Sur la Figure 3.12 un seul observateur indique les marches, la rampe et le mur.



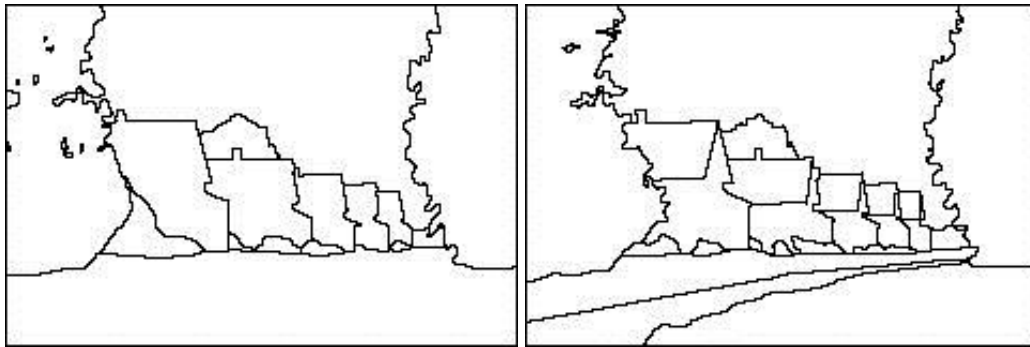


Figure 3.11 : Images de BSDS300, de gauche à droite et de haut en bas: l'image originale en couleur, la fusion des cartes de contours des observateurs, la carte de contours d'un observateur et la carte de contours d'un deuxième observateur.

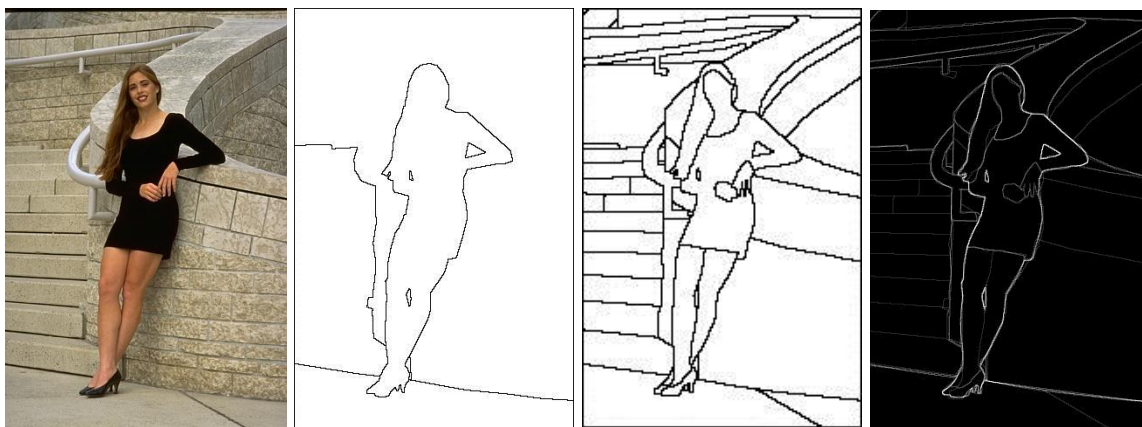


Figure 3.12 : Images de BSDS300, de gauche à droite : l'image originale en couleur, la carte de contours d'un observateur, la carte de contours d'un deuxième observateur, et la fusion des cartes de contours des observateurs

La banque d'image BSDS300, n'est donc pas adaptée à notre besoin. Nous avons besoin d'une banque d'image proposant des vérités terrain liées à l'action, la tâche en cours par la personne observant la scène. En d'autres termes pour aider les malvoyants nous avons besoin de détecter les contours importants dans le contexte de l'image, de la situation. Pour une même image un même malvoyant ne va pas souhaiter les mêmes contours s'il doit trouver son chemin, trouver une maison, monter des escaliers ou reconnaître une personne, cf. Figure 3.11 et Figure 3.12.

3.2.1.2. Notre banque de test

Au vu du constat de la partie 3.2.1.1, j'ai donc réalisé une banque d'images, afin de pouvoir évaluer les détecteurs de contours de manière objective pour les personnes malvoyantes. La vérité terrain, carte de contours dessinés par des observateurs, n'étant pas unique mais

dépendante du contexte dans lequel on se place pour analyser l'image. Pour une même image, on peut en effet, dessiner différentes cartes de contours en fonction de la tâche que l'on se fixe. Les malvoyants n'ont pas tous la même perception d'une même image, du fait de leur déficience, il n'aurait pas été possible d'obtenir une seule carte de contours. De plus, réaliser, la collecte de ces cartes de contours est beaucoup plus compliquée auprès de personnes malvoyantes. Il n'est donc pas possible d'obtenir une vérité terrain ou des vérités terrain représentatives des besoins des malvoyants, nous sommes donc contraints de réaliser une évaluation réelle auprès de malvoyants de nos développements afin de valider les résultats obtenus.

Notre banque d'images a un autre objectif, à savoir représenter les situations que les malvoyants peuvent rencontrer au cours d'une journée, en disposant du point de vue d'une caméra miniature sur une paire de lunettes. J'ai utilisé une caméra miniature comprend un capteur CMOS rolling-shutter (présenté au chapitre 5) couleur analogique PAL. Cette caméra est équipée d'un objectif intégré de focale 3,1mm et d'ouverture $F=3,4$. Cette caméra était connectée à un convertisseur vidéo USB permettant de convertir le flux de cette caméra pour l'ordinateur en flux vidéo couleur de résolution 640 X 480 pixels.

Les images ont été réalisées (quelques exemples en Figure 3.13) pour répondre à plusieurs critères :

- être représentatives de situations rencontrées dans les tâches de la vie quotidienne telles que détecter les marches d'un escalier, marcher dans un couloir peu ou mal éclairé, éviter des quilles de couleurs foncées sur un trottoir,...
- être représentatives des situations d'éclairement rencontrées par les patients : ombres, sur-illumination, sous-illumination, zone saturée
- prendre en compte le point de vue du patient utilisant une paire de lunettes de réalité augmentée (dispositif présenté au chapitre 5) : en effet, la caméra, située dans la paire de lunettes subira les mouvements de la tête, aura un point de vue légèrement plongeant sur la scène. Ce flux vidéo issu de cette caméra contiendra certains artefacts dus à sa petite taille, à son électronique et à son optique restreinte.



Figure 3.13 : Quelques images de la banque, la banque d'image complète est en annexe 2

Une fois la banque d'image réalisée, j'ai dû extraire les contours de manière automatique. Un grand nombre de travaux récents sur les détecteurs de contours reviennent souvent aux méthodes classiques telles que Sobel, et Canny. Il n'existe pas d'étude des détecteurs de contours pour les malvoyants. Je vais donc commencer par présenter les méthodes de Sobel et Canny puis une méthode récente de Canny Couleurs.

3.2.2. Sobel

Le détecteur de Sobel est un détecteur de contours basé sur le calcul de deux gradients sur une image pixel à pixel puis le calcul d'une carte de magnitude de ces gradients. Chaque gradient indique la direction de la plus forte variation du clair au sombre, ainsi que la variation locale de l'image dans cette direction ; celle du gradient.

Le détecteur de Sobel utilise deux masques de filtrage séparables :

$$\textit{Gradient horizontal} = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * \textit{Image originale}$$

$$\textit{Gradient vertical} = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \textit{Image originale}$$

Il est ensuite possible de calculer la magnitude de l'image de contours par :

$$\textit{Magnitude} = \sqrt{\textit{Gradient horizontal}^2 + \textit{Gradient vertical}^2}$$

L'orientation O du gradient est définie par :

$$\tan O = \frac{\textit{Gradient vertical}}{\textit{Gradient horizontal}}$$

Soit :

$$O = \tan^{-1} \frac{\textit{Gradient vertical}}{\textit{Gradient horizontal}}$$

Une fois la carte de magnitude calculée, il est possible en appliquant un seuil, valeur au-dessus de laquelle les pixels (points de l'image) sont considérés comme un contour, d'obtenir une carte de contours. Ce détecteur de contours est simple à calculer, les masques de convolution sont séparables ; cela permet à ce détecteur de nécessiter un temps de calcul court par rapport aux autres méthodes.

Afin d'adapter ce détecteur, il est possible de choisir différentes tailles de masques de convolution ou différentes valeurs de seuil.

Plus le masque de convolution est grand, plus il faudra de temps pour réaliser le calcul complet. Un plus grand filtre permet de prendre en compte plus de points voisins afin de définir si un pixel doit être considéré en tant que contour. Cela est surtout intéressant en cas d'image fortement bruitée, dans ce cas, il peut être intéressant d'appliquer un pré-filtrage gaussien, convolution avec un masque gaussien, afin d'atténuer le bruit.

Le choix de la valeur seuil permet de définir les pixels à considérer, en tant que contours, par rapport à leur valeur de magnitude. Cela est valable uniquement si les pixels à considérer en tant que contours ont une magnitude plus élevée que les pixels à ne pas considérer. Seule la valeur de la magnitude du pixel est prise en compte pour définir ce pixel en contour, il n'existe aucune prise en compte du voisinage de chaque pixel dans cette décision, cela pourrait être ajouté. La prise en compte de l'orientation de la magnitude et du voisinage de

chaque pixel pourrait permettre d'obtenir des résultats proches de ceux du détecteur de Canny.

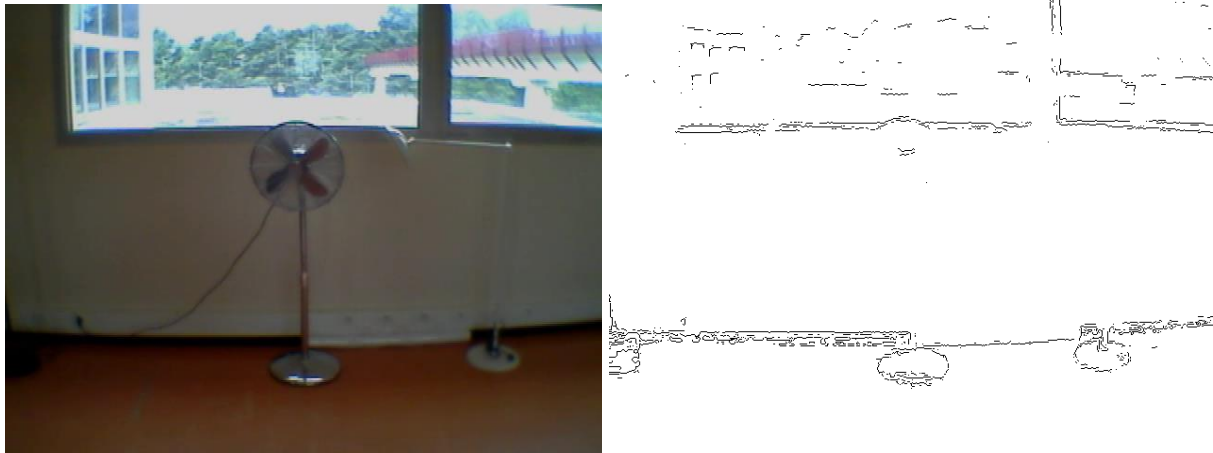


Figure 3.14 : (à gauche) Image originale ; (à droite) image filtrée par le filtre de Sobel et seuillée (valeur 100).

La Figure 3.14 présente une image source et le résultat seuillé du filtre de Sobel dans le but de conserver uniquement les pixels importants de la carte de magnitude. En effet, pour aider les malvoyants, il ne faut afficher que les contours les plus importants ; lorsque l'on affiche des informations sur un système à verres transparents une partie de la réalité est masquée. Lorsque l'on affiche des images complètes, la réalité n'apparaît que très légèrement. Dans le cas des malvoyants, il est important de laisser apparaître la réalité et d'afficher uniquement les contours importants. Pour le détecteur de Sobel, la carte de magnitude est uniquement seuillée, il n'y a pas d'analyse locale pour conserver les pixels présentant la magnitude la plus élevée localement, ou de lien entre les pixels pour lier les contours et conserver tout un contour. Ce détecteur propose souvent trop de contours, comme sur la Figure 3.14, pour la zone de la fenêtre, mais trop peu de contours pour détecter la lampe et le ventilateur.

3.2.3. Canny

Pour répondre aux besoins de détection de contours, John Canny a proposé un algorithme capable de détecter dans certaines conditions les contours de manière efficace. Pour cela, il a tout d'abord défini les critères définissant un contour

- sa localisation : la distance entre le contour réel et le contour détecté doit la plus faible possible

- son rapport signal sur bruit élevé : cela permet d'être sûr que le contour détecté est bien un contour réel. Le détecteur se doit de détecter tous les contours sans détecter de faux contours.
- Son unicité : un contour ne doit pas solliciter plusieurs pixels contours, il faut détecter et conserver uniquement les pixels contours les plus importants. Il doit exister un seul pixel contour pour chaque zone du contour réel pas une multitude de pixels.

Ces trois critères sont toujours valables dans le cas de l'assistance visuelle aux malvoyants. En effet, la position des contours affichés doit correspondre à la position réelle des contours (localisation). De plus, afin de ne pas saturer le malvoyant d'informations superflues, le contour détecté doit être unique (unicité). Enfin, le contour affiché doit être le plus pertinent possible donc son rapport signal sur bruit doit être le plus élevé possible. Les critères choisis par Canny pour définir ses contours correspondent aux besoins des malvoyants en termes de contours.

Afin de réaliser ce détecteur de contours, il a été défini le rapport signal sur bruit SNR et une fonction de localisation :

$$SNR = \frac{A \left| \int_{-w}^0 f(x) dx \right|}{\sqrt{\int_{-w}^w f^2(x) dx}}$$

$$Localisation = \frac{A|f(0)|}{\sqrt{\int_{-w}^w f^2 dx}}$$

Il faut préfiltrer l'image, appliquer une dérivée de gaussienne, calculer une carte de magnitude et d'orientation, et enfin réaliser la suppression des maximums couplés à un seuillage par hystérésis pour conserver uniquement les vrais contours [38, 37, 130].

Il est tout d'abord appliqué à l'image un filtre gaussien qui permet de rendre le détecteur moins sensible au bruit. Ce filtrage est généralement réalisé avec un masque horizontal et donne une image A_x . De même, il est réalisé une image A_y avec un masque gaussien vertical. Ces deux images sont ensuite convoluées à des dérivées de gaussienne respectivement suivant leurs directions horizontales et verticales. On obtient alors les images G_x et G_y . Une fois ces images obtenues, nous pouvons calculer l'orientation et la magnitude des contours comme cela est réalisé pour le filtre de Sobel.

Afin de respecter le critère d'unicité des contours, il faut supprimer les non-maxima locaux. Ce filtrage est souvent réalisé simultanément au seuillage par hystérésis. Pour cela il faut

définir les seuils hauts et bas. Les pixels ayant une valeur supérieure au seuil bas sont considérés comme des contours potentiels. Si le contour potentiel a au moins un pixel dont la valeur est supérieure au seuil haut alors il est considéré comme un contour. Cette analyse est réalisée en fonction de chaque orientation pour 4 valeurs : 0° , 45° , 90° et 135° . Un contour ne doit contenir que les pixels les plus caractéristiques de ce contour. Cette partie est un point important du détecteur de Canny comme la convolution avec la dérivée de gaussienne.

Ce détecteur dispose de trois paramètres qui permettent de l'adapter aux images : la taille du filtre gaussien, les seuils haut et bas permettant de l'adapter aux images, cf. Figure 3.15.

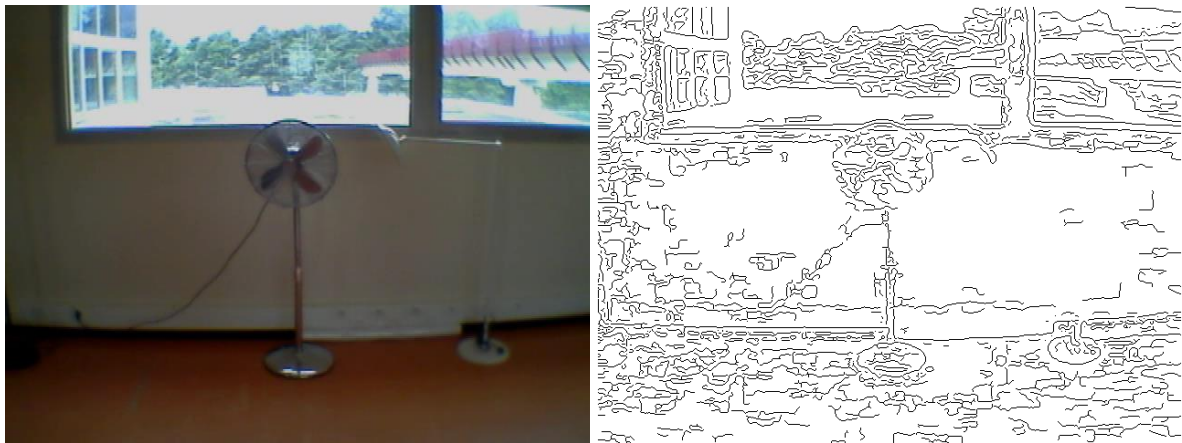


Figure 3.15 : (à gauche) Image originale ; (à droite) image filtrée par le filtre de Canny

La Figure 3.15 représente une image source et le résultat du détecteur de Canny pour un sigma égale à 3, filtre de taille 3x3, un seuil bas de 50 et un seuil haut de 100.

Sur cette figure, Beaucoup de contours sont détectés. Ce détecteur nous fournit beaucoup plus de contours que le filtre de Sobel. Les paramètres seuils bas et haut permettent d'adapter la quantité de contours en fonction de la quantité souhaitée. Afin de réduire la quantité de détails sur la carte de contours, il faudrait augmenter le paramètre sigma du filtre gaussien afin de filtrer davantage l'image.

Le filtre gaussien permet de réduire le bruit présent dans l'image, avec une trop faible valeur le détecteur donne trop de contours en sortie, avec une valeur trop élevée, les informations contenues dans l'image sont lissées, il sera donc détecté peu de contours.

Les paramètres influencent la qualité de la carte de contours obtenue. De même certaines implémentations du détecteur de contours de Canny, le filtre gaussien et les dérivées de gaussiennes sont remplacés par des masques proches du filtre de Sobel, pour plus de simplicité de code et la rapidité d'exécution sur ordinateur [130, 178], ce qui peut impacter la qualité de la carte de contours.

3.2.4. Canny couleur

Le détecteur de Canny fonctionne sur une image en niveau de gris uniquement. Cependant, les images sont souvent en couleurs, il est intéressant d'analyser si la couleur peut apporter une amélioration pour la détection de contours.

Certaines équipes de recherche ont proposé des détections de contours, ou des seuillages en utilisant des descripteurs couleurs afin de les rendre invariantes à la luminance, ces méthodes donnent de bons résultats dans le cas d'images comprenant des couleurs différentes pour chaque élément de la scène [69, 68, 188, 111].

Il a notamment été développé une adaptation du détecteur de Canny pour les images couleur « Canny couleur » [185]. Ce détecteur de contours est présenté ici, car il prend en compte l'information de couleur. Nous l'avons utilisé lors de l'évaluation avec les malvoyants du chapitre 4, en tant que détecteur de contours récent (2006-2007) et prenant en compte l'information de couleur. Ce détecteur de contours basé sur la couleur peut utiliser différents espaces couleurs. Il se base sur la conversion de l'image couleur en tenseur couleur constant (Color Constancy Tensor). Il est possible de choisir plusieurs tenseurs différents en fonction de l'image et du résultat souhaité. Une des difficultés dans l'usage de cette méthode est le choix du tenseur couleur ; certains sont quasi invariants, d'autres invariants aux ombres ou aux sur-illuminations. Une fois le tenseur calculé, il est calculé les dérivées suivant les directions horizontales et verticales dans l'image. Il est ensuite possible d'appliquer la fin de l'algorithme de Canny à savoir le calcul de la magnitude, des orientations. Il est enfin réalisé le seuillage à hystérésis et la suppression des non-maxima, cf. Figure 3.17. Ce détecteur est donc proche du détecteur de Canny, mais en se basant sur le calcul du tenseur couleur de l'image au lieu d'une image en niveau de gris.

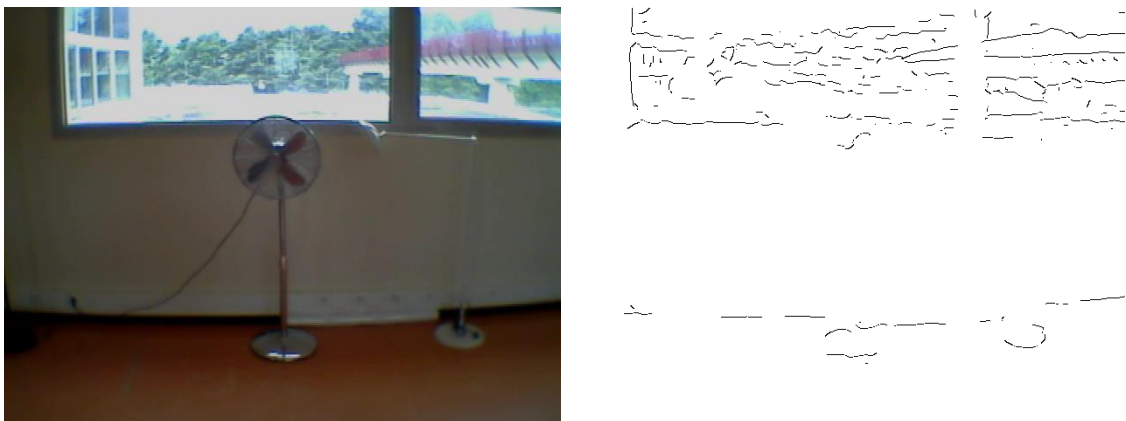


Figure 3.16 : (à gauche) Image originale ; (à droite) image filtrée par le filtre de Canny Couleur

Pour le résultat de la Figure 3.16, le tenseur utilisé est le tenseur invariant aux reflets et ombres (specular shadow-shading quasi invariant & variant). Ce détecteur donne des résultats intéressants ; lorsque les éléments ont suffisamment de variations de luminance ou de couleur, mais ne permet pas de marquer toute la délimitation entre le mur et le sol par exemple.

3.2.5. L'aspect multi-échelle : les pyramides

Comme cela est visible avec le détecteur de Canny, la variation du sigma du filtre gaussien permet de détecter soit tous les contours soit uniquement les contours nous intéressant. Dans une image, il existe plusieurs niveaux d'informations. Afin d'analyser chaque niveau, il faut une intervention sur le sigma. Cela n'est pas réalisable par un malvoyant. Il serait utile de disposer d'une solution d'analyse multi-échelle.

Lorsque l'on ne connaît pas la taille à priori des éléments à détecter, la taille des contours à extraire, il peut être intéressant de réaliser un même traitement en prenant en considération plusieurs fréquences spatiales ou plusieurs tailles de l'image. Lorsque l'on modifie uniquement la taille de l'image ou que l'on prend en compte plusieurs fréquences spatiales, il est possible de réaliser un détecteur multi-échelle prenant en compte plusieurs niveaux d'informations. Lorsque l'on réalise les deux en même temps, à savoir la modification de la taille de l'image et les différentes fréquences spatiales, on prend alors en compte le maximum de niveaux d'information.

Il est très difficile de connaître la taille des éléments à détecter, la fréquence spatiale à analyser pour obtenir les informations souhaitées et le niveau de détails souhaité. Pour une même image, il est possible d'obtenir trop peu d'informations, si on analyse cette image dans un niveau de détail trop faible, trop d'informations, si on analyse cette image dans un niveau de détail trop élevé. Comme le montre la Figure 3.17, certains niveaux de détails dans les cartes de contours rendent leur exploitation difficile, soit par l'absence d'un grand nombre de contours importants, soit par la présence de beaucoup trop de contours et masquant ainsi les informations importantes.

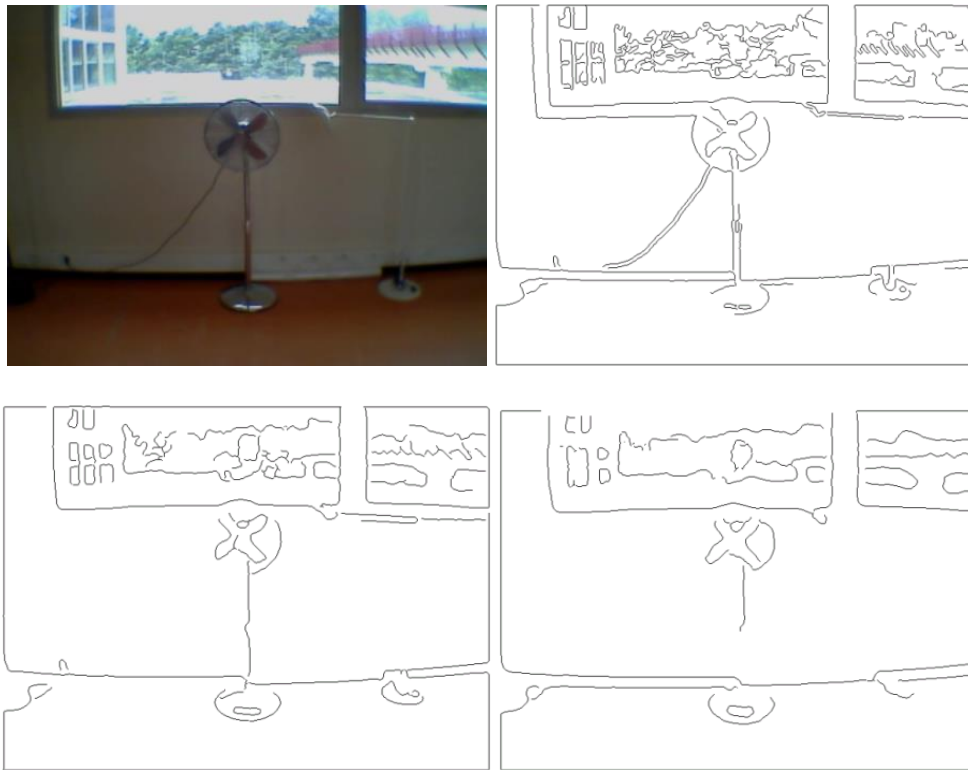


Figure 3.17 : Filtrage à différentes échelles d'une même image (de gauche à droite et de haut en bas) l'image d'origine, l'image avec un niveau d'information élevé (0.5), l'image avec un niveau d'information acceptable (3.5) et l'image avec un niveau d'information faible (5.5).

Afin de réaliser cette analyse sur différentes tailles d'images, il faut utiliser une pyramide. Il s'agit de représenter la même image à chaque niveau en réduisant sa taille et en lui appliquant un filtre généralement un filtre gaussien. Comme le montre la Figure 3.18, le premier niveau, la base, a la plus grande résolution et également les détails les plus fins. Le dernier niveau ayant la plus petite résolution, contient les informations les plus simples.



Figure 3.18 : Représentation d'une pyramide⁴⁹

⁴⁹ [http://fr.wikipedia.org/wiki/Pyramide_\(traitement_d'image\)](http://fr.wikipedia.org/wiki/Pyramide_(traitement_d'image))

En appliquant un même traitement d'images à tous les niveaux de la pyramide puis en fusionnant, synthétisant ces résultats, il est possible de collecter les informations les plus pertinentes de l'image. Cette technique de la pyramide permet de naviguer rapidement entre chaque niveau d'information. Dans une pyramide, il est facilement possible d'accéder aux voisinages d'un point dans les niveaux d'informations différents. Il existe un voisinage hiérarchique : un pixel d'un niveau est voisin des points qui sont soit ses voisins dans un autre niveau de la pyramide soit confondus.

L'utilisation de pyramide permet d'analyser une même image avec plusieurs échelles afin de pouvoir s'adapter aux différentes dimensions, fréquences des informations de l'image [116, 56]. Les pyramides permettent en traitement d'image de fusionner une grande quantité d'informations [41], d'améliorer les performances de certains algorithmes [177, 154, 94], ou de visualiser les données [89, 23, 78] ou détecter les formes souhaitées [105].

La pyramide, permet d'analyser un certain nombre de fréquences spatiales plus rapidement et surtout d'obtenir des résultats pertinents sans avoir besoin de définir à l'avance la fréquence spatiale optimale de l'image à analyser. Certes, cette méthode nécessite un temps de calcul plus long, car il faut générer tous les niveaux de la pyramide, appliquer le traitement sur tous les niveaux puis fusionner ces niveaux, mais cela permet d'adapter des outils et les rendre moins sensibles aux différents niveaux d'information.

3.2.6. Filtrage de contours par pyramide

En se basant sur les besoins des personnes malvoyantes, sur nos évaluations, nos tests des méthodes de l'état de l'art, la détection de contours peut apporter une aide aux personnes malvoyantes. Pour cela, il est nécessaire de disposer d'un détecteur adapté. Comme nous venons de le voir le filtre de Sobel ou celui de Canny ne sont pas toujours adaptés en fonction du niveau d'information requis. Un système d'analyse multi-échelle peut permettre de prendre en compte les différents niveaux d'informations. En fusionnant ces méthodes, j'ai obtenu une nouvelle méthode. Cette nouvelle méthode filtre les contours sur une base de pyramide. Il est nécessaire de réaliser 3 grandes étapes, cf. Figure 3.19 :

- 1 un filtrage ajustable en temps réel : afin de réaliser le pré-filtrage du bruit et le réglage de la quantité de contours
- 2 une détection de contours sur plusieurs niveaux d'une pyramide : cela permet de prendre en compte tous les niveaux d'information lors de la détection de contours

- 3 une fusion de ces cartes, la sélection des contours pertinents et leur restitution sur une carte unique de contours : cela permet de conserver uniquement les contours les plus importants pour comprendre la scène.



Figure 3.19 : Schéma général de l'algorithme

3.2.6.1. Filtrage ajustable

L'image d'origine est une image couleur, cette image est convertie en image à niveaux de gris.

Pour convertir les images RGB en image à niveau de gris, il est possible de calculer la luminance, de convertir l'image dans un espace couleur différent tel que HSV et d'extraire uniquement le plan de luminance. L'espace couleur HSV est très intéressant car il sépare les informations couleurs en fonction de son plan achromatique [42], sans prendre en compte l'information de couleur de l'image. J'ai donc choisi le plan de la « valeur » de l'espace couleur HSV, car il me permet d'obtenir les meilleurs résultats, sans être fortement impacté par les zones saturées des images. Lors de l'utilisation de l'information de teinte ou de couleur, il est détecté certaines zones saturées, et les aberrations liées au capteur, cf. Figure 3.20.

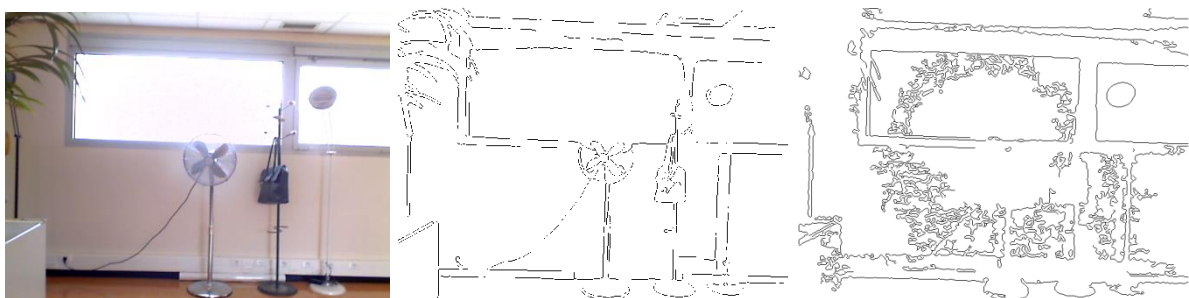


Figure 3.20 : De gauche à droite : image originale, carte de contours sur la luminance (Value HSV) et carte de contours sur la Teinte (Hue HSV)

L'image de luminance ainsi obtenue peut maintenant être filtrée par un filtre gaussien. Le filtre gaussien est connu pour conserver les propriétés des informations contenues dans

l'image [34, 13]. Ce filtre passe-bas permet d'une part, d'atténuer le bruit de l'image à analyser et d'autre part d'ajuster la sensibilité du détecteur de contours global. Comme expliqué précédemment, un besoin important des malvoyants est d'adapter la quantité d'information en fonction de la situation. Le résultat final de cette méthode dépendra de la taille de ce filtre gaussien, cf. Figure 3.21.

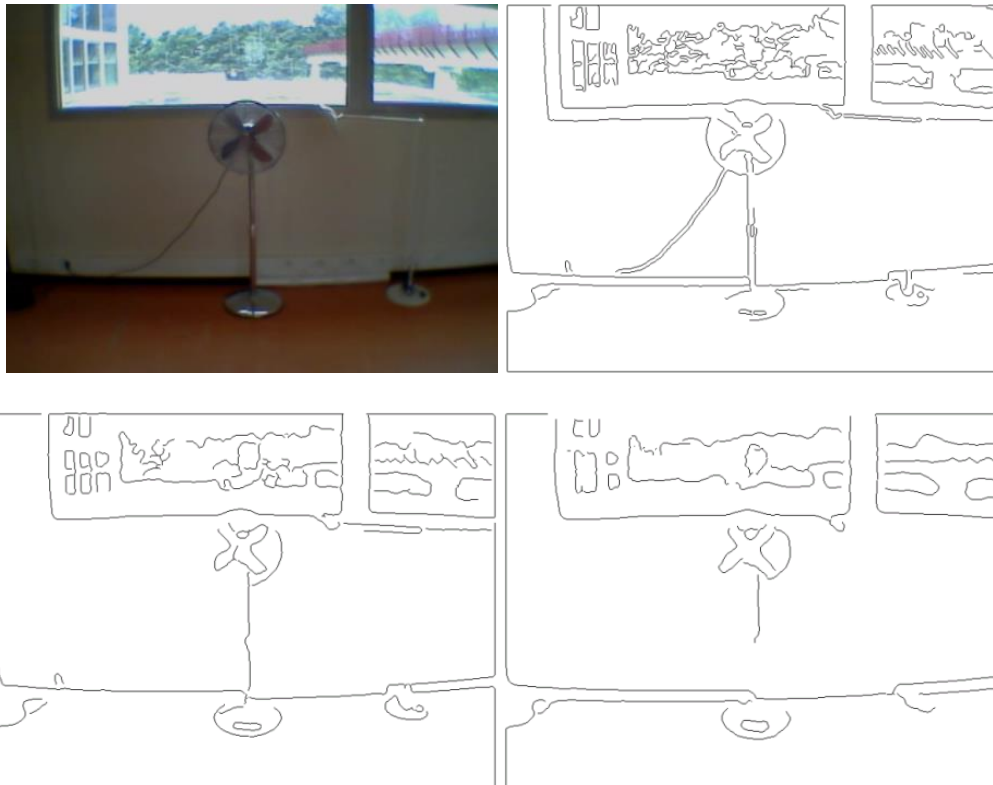


Figure 3.21 : Exemples de cartes de contours pour une même image : (de gauche à droite et de haut en bas) ; image originale, images pour un filtrage gaussien différent sigma de 0.5, 3.5 et 5.5.

3.2.6.2. Détection de contours multi-échelle

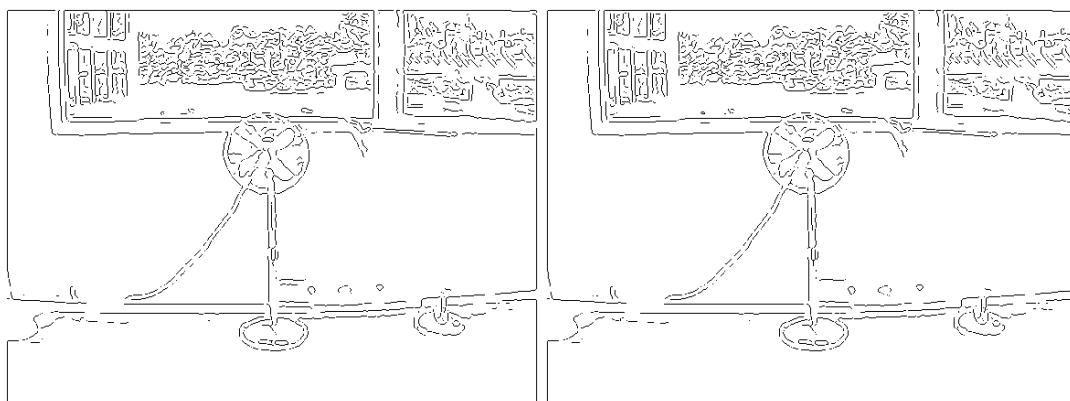
Maintenant que l'image a été prétraitée, nous pouvons générer une pyramide sur la base d'une méthode adaptée de Burt and Adelson [34]. Notre implémentation combine à la fois un filtrage passe-bas, et une réduction de la taille de l'image pour chaque niveau de la pyramide.

Nous obtenons ainsi une série d'images. Les deux premiers niveaux de la pyramide ont la même taille seul le filtrage gaussien diffère cela permet de conserver certaines informations qui seraient altérées par la combinaison du filtrage gaussien et la réduction de la taille de l'image. La plus petite carte de la pyramide, le plus haut niveau de celle-ci, correspond aux fréquences spatiales les plus faibles. La plus grande carte correspond aux informations d'origine convoluées uniquement au masque gaussien.

Nous avons analysé l'impact sur la carte de contours du nombre de niveau de la pyramide et réalisé une validation du nombre de niveau de manière expérimentale. Avec 1, 2 ou 3 niveaux, il reste encore beaucoup trop de petits contours. À partir de 4 niveaux, les cartes de contours sont exploitables par les malvoyants. À partir de 5 niveaux, nous commençons à altérer les contours importants. La pyramide utilisée dans cette méthode est constituée de 4 niveaux ; le quatrième niveau a une taille 160 X 120 pour une image d'origine de taille 640 X 480.

La différence entre deux niveaux de la pyramide peut également donner des contours. Dans notre cas, sur nos images de test, cela donnait des contours mais également beaucoup de faux contours liés aux bruits résiduels dans les premiers niveaux (ceux de plus grandes résolutions). De plus, la soustraction de deux niveaux donne des contours qui ne sont pas filtrés. Le détecteur de Canny filtre les contours par hystérésis et/ou suppression des non-maximas locaux (des contours peu marqués). Ces méthodes de filtrage nécessitent de connaître l'amplitude et l'orientation du gradient ; ces valeurs ne sont pas accessibles immédiatement dans le cas de la soustraction de niveaux de la pyramide.

Chaque carte de niveau de cette pyramide est ensuite analysée afin de détecter tous les contours de celles-ci, cf. Figure 3.22. Il est possible d'utiliser la méthode de détection de son choix par exemple Sobel, Canny ... Comme la pyramide est dérivée du filtre gaussien, nous pouvons considérer que le bruit est suffisamment filtré par l'usage de notre méthode, un pré-filtrage avant la détection de contours n'est plus nécessaire. Dans le cas du détecteur de contours de Canny, il est nécessaire de définir la taille du filtre, les valeurs minimale et maximale pour le seuillage par hystérésis. Il est possible de calculer les seuils haut et bas à partir de l'histogramme cumulé de la carte de magnitude : le seuil haut vaut la valeur la plus proche des 70 % de cet histogramme, le seuil bas vaut 40 % du seuil haut.



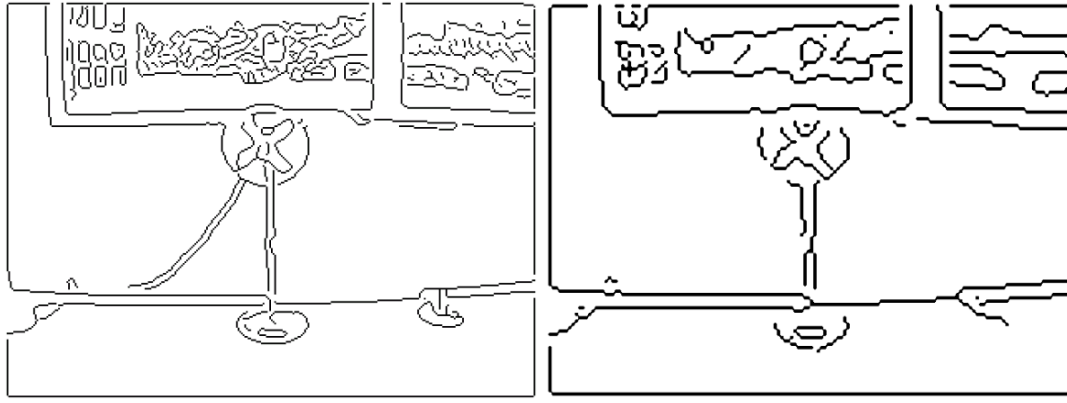


Figure 3.22 : Cartes de contours pour une même image à chaque niveau de la pyramide

3.2.6.3. Fusion, sélection et restitution des contours

Tous les niveaux de la pyramide sont redimensionnés à la taille de l'image d'origine. Les cartes de contours remises à la taille d'origine sont combinées entre elles par la relation ci-dessous :

$$C(x, y) = \sum_{i=1}^H 2^{H-i} \cdot S_i(x, y) \quad (1)$$

Où H est le niveau de la pyramide, S_i correspond à la carte de contours binaire redimensionnée à la taille d'origine du niveau i de la pyramide et C : la carte de contours combinés.

La carte de plus haute résolution est pour $i=1$. Cette carte de contours est celle présentant les informations les plus fines. La carte de niveaux H contient de plus gros blocs, donc les informations plus éloignée de la réalité et bénéficie d'un coefficient plus faible.

Dans nom cas, 4 niveaux de la pyramide sont utilisé. A partir du 5^{ème} niveau trop peu de contours sont pris en compte. L'analyse réalisée est faite sur un niveau d'information trop éloigné de l'image de départ. De même, avec uniquement 3 niveaux le filtrage n'est pas suffisamment marqué.

Pour un faible cout calculatoire et pour plus de simplicité, l'approche par addition pondérée des cartes présente l'avantage de ne pas nécessiter de paramètre tel que la taille des images ou de forme d'éléments structurants, ou encore la taille entre les cartes consécutives comme cela peut être le cas dans la littérature [129, 104].

Comme expliqué par la relation (1), nous avons pondéré les cartes de contours suivant cette relation, puis nous avons réalisé la somme pour chaque pixel pour tous les niveaux de la pyramide, cf. Figure 3.23.

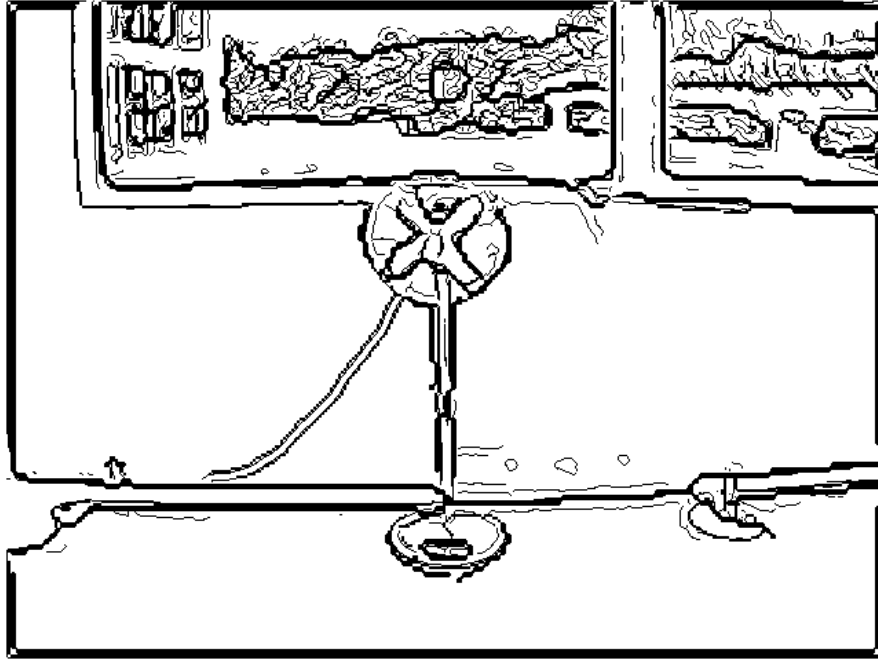


Figure 3.23 : Image de fusion des cartes de contours

Nous obtenons ainsi la carte de contours, il nous reste maintenant à détecter les pixels contours importants permettant de conserver uniquement les contours pertinents, cf. Figure 3.24. Pour cela nous réalisons la détection des graines par la méthode (2) :

$$C_b(x, y) = \begin{cases} 1 & \text{if } C(x, y) > \text{une valeur prédéfinie} \\ 0 & \text{sinon} \end{cases} \quad (2)$$

Cette technique permet de conserver la localisation des contours issus de la carte 1, la plus grande, tout en bénéficiant de la sélection des contours ci-dessus. Avec cette approche cumulative, les contours avec la plus grande amplitude sont ceux présents dans toutes les cartes. Nous utilisons les résultats binaires (0 ou 1) des cartes de contours de chaque niveau de la pyramide car à ce stade nous souhaitons détecter les contours présents dans toutes les cartes. La valeur maximale possible pour un contour est définie par :

$$\text{Maximum } C = \sum_{i=1}^H 2^{H-i} = 2^H - 1 \quad (3)$$

Si le seuil des contours est compris entre la valeur C maximale et 2^{H-1} alors les contours sont au moins présents dans la carte la plus basse. Le choix de cette valeur permet de sélectionner soit les contours importants uniquement soit de faire apparaître les contours moins marqués ou les contours présents dans les textures ; ces contours sont surtout présents dans la carte basse de la pyramide.

Dans le cas où le critère est choisi égal au maximum de C , alors cela revient à faire un ET logique entre tous les niveaux de la pyramide, chaque contour doit être présent dans toutes les cartes de contours pour être conservé.

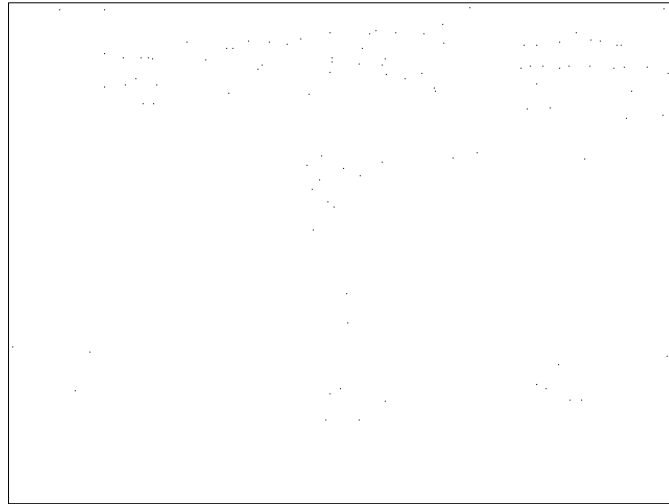


Figure 3.24 : Carte représentant les graines sélectionnées

La dernière étape de notre méthode consiste à conserver uniquement les contours de la carte basse de la pyramide pour lesquels nous avons une graine. Pour cela, la carte basse (de plus grande résolution) est labellisée (il est attribué une étiquette unique à chaque contours), puis nous analysons pour chaque graine à quel label de contours, elles appartiennent. Une fois la liste des labels de contours à conserver réalisée, nous pouvons dessiner la carte de contours finale, cf. Figure 3.25. Seuls les contours issus de la première carte (celle de même résolution que l'image d'origine) sont dessinés, cela permet de conserver les propriétés du détecteur de Canny (unicité, localisation et rapport sur bruit) sans dessiner les contours peu importants.

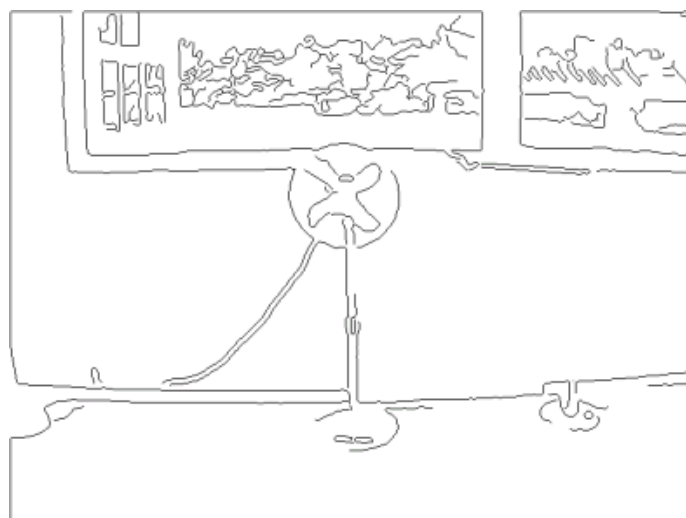


Figure 3.25 : Carte de contours finale

Cette méthode permet de corriger des problèmes éventuels liés soit au décalage des contours dans l'image, soit à l'altération de la suppression des non-maxima locaux présente dans certaines méthodes telles que le détecteur de Canny [49]. Cette dernière étape permet de conserver le critère de localisation des contours.

Nous avons donc maintenant une méthode permettant de détecter des contours et de pouvoir adapter la quantité de contours en adaptant uniquement un seul paramètre (la taille du filtre gaussien).

3.2.6.4. Présentation de résultats

Dans cette partie, je compare mon outil par rapport à l'état de l'art de manière qualitative. Pour cette analyse les mêmes paramètres pour chaque détecteur ont été utilisés pour toutes les images. Si le détecteur dispose d'un mode de calcul automatique des seuils, celui-ci a été utilisé ; Etant donné qu'il a été utilisé les implémentations de Matlab pour Sobel et Canny, le filtre de Canny a bénéficié du calcul automatique de ses seuils par Matlab⁵⁰. Pour le Canny Couleur j'ai utilisé le code de van de Meijer [185].



Figure 3.26 : De gauche à droite et de haut en bas : image originale, et les résultats pour le filtre de Sobel, le filtre de Canny, le filtre Canny couleur et notre pyramide

Sur la Figure 3.26, notre méthode présente des contours plus faciles à interpréter que les résultats des autres méthodes, surtout pour le passage piéton. En effet, les résultats des détecteurs de Sobel et Canny couleurs sont incomplets, en particulier pour le passage piéton,

⁵⁰ <http://www.mathworks.fr/>

et fractionnés. Le détecteur de Canny donne lui trop de contours et sature même les informations pouvant être intéressantes, les contours utiles sont masqués par la quantité de contours détectés.

Sur la Figure 3.27, le filtre de Sobel détecte trop peu de contours pour aider un malvoyant à identifier le passage qu'il peut emprunter. La méthode pyramidale permet de signaler la zone de déambulation et la présence de murs sur les côtés par une densité de contours différente.



Figure 3.27 : Image originale et les images pour le filtre de Sobel et la méthode pyramidale



Figure 3.28 : Image originale et les images pour le filtre de Canny et méthode pyramidale

Sur la Figure 3.28, le résultat du filtre de Canny est nettement saturé d'informations, ce filtre donne beaucoup trop d'informations. La méthode pyramidale se base sur le même type de carte de contours mais utilise d'une part, un pré-filtrage permettant de réduire le bruit de l'image avant la détection des contours et d'autre part, la partie pyramidale pour sélectionner les contours à conserver et ceux de moindre importance à supprimer. Ces deux aspects de mon détecteur sont importants pour cette image : le filtrage permet de détecter des contours moins fractionnés sur le bord des marches et sur la main courante, la partie pyramidale permet de supprimer tous les contours issus des variations de luminance dans l'image dues aux textures sur les marches et le sol.

Le résultat du filtre Canny-couleur de Gevers et Van de Meijer donne de bons résultats pour certaines images et des résultats quasi-inexploitables pour d'autres images, cf. Figure 3.29, ici le micro-onde est très mal détecté, tout comme le bord du réfrigérateur et les torchons. Il est très difficile de trouver un ensemble de paramètres pour ce filtre permettant de l'appliquer à

toute notre banque d'images sans avoir à adapter les paramètres. Il peut s'avérer difficile de trouver une série de paramètres permettant de répondre aux besoins des personnes malvoyantes.



Figure 3.29 : Image originale et les images pour le filtre de Canny-couleur et notre pyramide

3.2.6.5. Schéma général de l'algorithme

Nous présentons maintenant un schéma global détaillé, cf. Figure 3.30, de notre méthode pour tout détecteur de contours, et pour un filtre gaussien à définir en fonction du résultat souhaité.

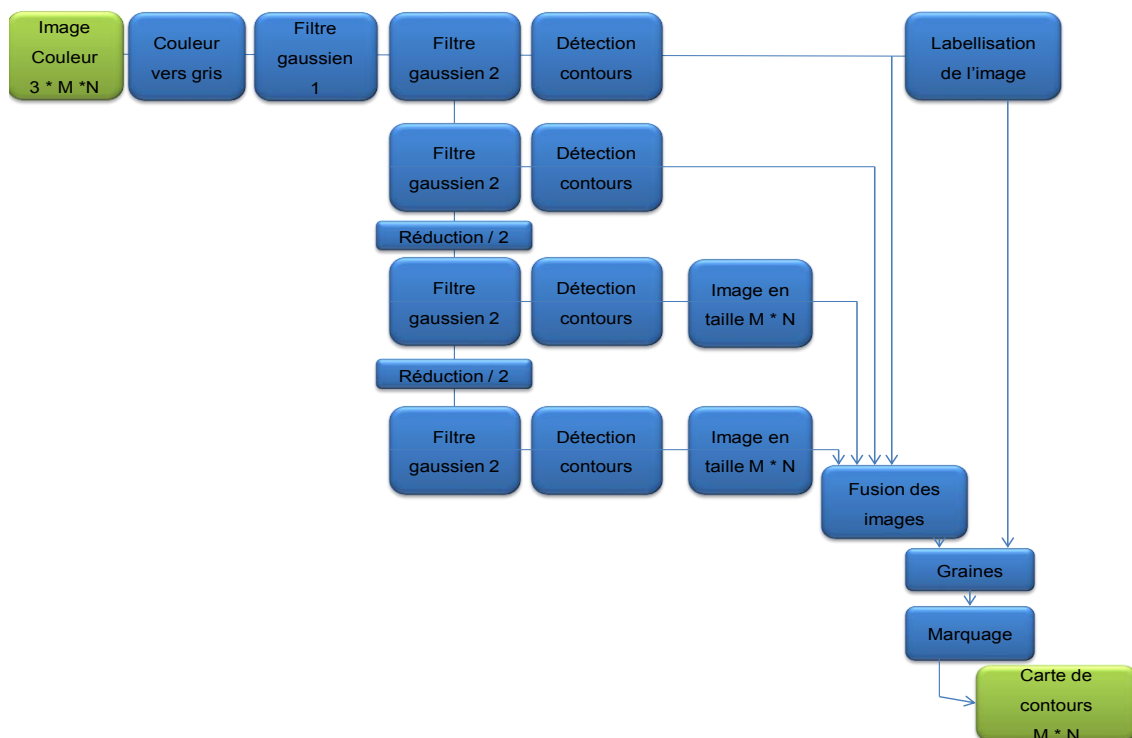


Figure 3.30 : Schéma détaillé de la méthode

Dans le cas qui nous intéresse, à savoir une méthode pour détecter et régler la quantité de contours répondant aux besoins des malvoyants, nous avons utilisé une pyramide à 4 niveaux (cf. section 3.2.6.2). Le détecteur de contours est le détecteur de Canny et le filtre gaussien 2

est un filtre gaussien standard de taille 3 X 3, sigma 0,5. Le plus petit niveau de notre pyramide a des images de taille 120 X 160 pixels. Le niveau de filtrage pour la fusion des cartes de contours de la pyramide est de 14, autrement dit les contours doivent être marqués dans toutes les cartes pour être retenus. On conserve uniquement les contours les plus marqués.

3.2.6.6. Comparaisons quantitatives aux méthodes de l'état de l'art

Un certain nombre de travaux concernant l'étude des détecteurs de contours ont été réalisés [127], ces études nécessitent de pouvoir évaluer les résultats de chaque détecteur de contours par rapport à une ou des références [118, 161, 91]. Certaines équipes ont mis en place des métriques permettant d'analyser des contours obtenus avec des images de contours références [40, 24]. Ces méthodes d'analyse nécessitent de connaître une image vérité terrain. Afin de comparer notre méthode aux méthodes de l'état de l'art sans vérité terrain, sans banque d'images objective existante, nous avons réalisé quelques cartes de contours. Nous avons pour cela dessiné les contours sur quelques images de la banque d'images ; j'ai pris en compte le contexte de mobilité, j'ai dessiné uniquement les éléments dans la zone de déambulation ou les éléments avec lesquels la personne peut se heurter, ou ceux très fortement marqués. Ce contexte de mobilité est le paramètre le plus important dans cette analyse, en effet pour évaluer des solutions pour aider les malvoyants, il est nécessaire de prendre en compte l'usage et le contexte dans lequel ces solutions sont utilisées. Ces images de tests ont été réalisées sur 4 images de la banque d'images seulement, cf. Figure 3.31, dans le seul but de réaliser une analyse de ces indicateurs malgré l'absence de banque d'images conséquente. Ce type d'analyse doit être réalisé sur un grand nombre d'images, ici, nous avons uniquement une tendance.

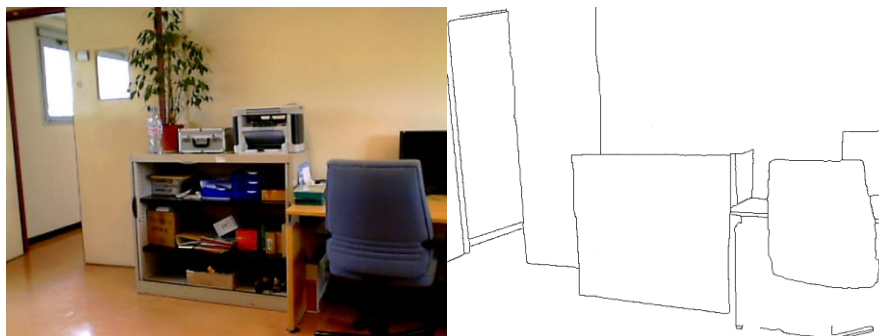


Figure 3.31 : Exemple d'image et sa vérité terrain, les 4 images sont en annexe 3

Il est nettement visible sur les courbes présentées en annexe 4, que ces indicateurs sont difficilement exploitables ici, car il nous faut réaliser ces analyses sur une série d'images plus conséquentes, et surtout sur une vérité terrain réalisée auprès d'un plus grand nombre de personnes. Idéalement, la vérité terrain devrait être réalisée avec des personnes malvoyantes. Cela compliquerait l'analyse des résultats, car nous n'aurions pas une, mais des vérités terrain.

Ces indicateurs permettent néanmoins de mettre en évidence un de nos objectifs pour le développement de notre méthode pyramidale à savoir l'adaptation progressive de la quantité de contours. Il est visible sur un grand nombre de courbes une progression des valeurs indiquant une diminution du nombre de contours détectés, plus le paramètre de filtrage est élevé. Tous les autres paramètres sont restés fixes, pour Canny, il aurait été possible de faire varier un ou plusieurs paramètres, mais le choix de celui-ci aurait eu une influence non-négligeable sur le comportement de ce filtre.

3.3. Conclusion

Peli a démontré que les contours apportent une aide pour les personnes malvoyantes, mais que les méthodes actuelles n'apportent pas entière satisfaction. De plus, un détecteur de contours idéal doit pouvoir s'adapter à la situation, aux besoins du patient en fonction de sa vision. Il peut être utile que la méthode s'adapte à la quantité de contours désirée par le patient. Celui-ci peut ainsi adapter le détecteur en fonction de la scène où il se trouve et en fonction de son besoin, de sa pathologie. Le contexte de chaque scène est très important pour détecter les contours les plus utiles aux malvoyants.

J'ai présenté certaines méthodes et mis en œuvre une méthode qui répond au besoin d'un détecteur adaptable par un patient déficient visuel. Cette méthode peut utiliser différents détecteurs de contours, mais se base sur une pyramide pour filtrer les contours pertinents à afficher à la personne.

Nous avons comparé quantitativement les différents extracteurs de contours. Il est nécessaire de comparer la quantité des contours fournie par chaque méthode ainsi que leur pertinence par rapport à une vérité terrain. Cette vérité terrain doit dépendre du contexte de chaque scène. Il a été mis en évidence la faible détection de contours par le filtre de Sobel et le peu de liens entre ce résultat et la vérité terrain. Le détecteur de Canny extrait généralement trop de contours, jusqu'à saturer les images dans certaines conditions. Le réglage du sigma ne permet

pas à lui seul de répondre favorablement au besoin des malvoyants, il reste d'abord beaucoup de contours, puis lorsque le sigma augmente, les contours ne sont plus tout à fait fidèles aux contours initiaux (différence de forme et de position). Notre méthode pyramide permet de détecter les contours en conservant la localisation des contours et surtout de sélectionner parmi les contours détectés uniquement les contours les plus importants (ceux présents dans plusieurs niveaux d'information).

Il a également été mis en évidence le filtrage réglable des contours en fonction des besoins de l'utilisateur par le réglage d'un seul paramètre. Ce paramètre a un impact direct sur la quantité de contours, leur localisation et la distance entre cette carte de contours et la position des « vrais contours ».

La vérité terrain utilisée ici dépend uniquement de quelques personnes, cependant pour bien évaluer notre détecteur de contours, il faudrait plusieurs vérités terrain en fonction des besoins de chaque patient. Cela permettrait de mettre en évidence l'apport des différents réglages possible du détecteur.

Nous avons comparé visuellement les résultats des différentes méthodes. Suite à cette analyse sur quelques images, il ressort que notre méthode peut apporter, à priori, un plus non-négligeable pour les patients malvoyants. Afin de valider cet apport, il faut réaliser un essai avec ces déficients.

4. Évaluation de l'intérêt du rehaussement des contours pour les malvoyants

Je viens de présenter le développement d'un détecteur de contours afin de répondre aux besoins des malvoyants. Afin de valider ce détecteur, je vais dans un premier temps, dans la section 4.1 de ce chapitre, faire un bref état de l'art des évaluations d'outils de traitement d'image pour l'assistance des malvoyants. Je vais, dans la section 4.2, présenter le protocole de l'expérimentation mise en place afin d'analyser l'impact et l'apport des contours pour les malvoyants. Cette expérimentation avait également pour objectif de quantifier l'apport de chaque méthode de détection de contours. Dans la section 4.3, je vais présenter les deux groupes de malvoyants ayant participé à cette expérimentation. Les résultats de cette expérimentation sont présentés dans la section 4.4. Ces résultats sont présentés par thème afin de répondre dans un premier temps de montrer l'apport des contours. Il est également analysé les méthodes les unes par rapport aux autres. Enfin, je vais faire une conclusion sur l'apport des contours et de ma méthode de détection de contours dans la section 4.5.

4.1. État de l'art

Le rehaussement de contours a été évalué par l'équipe de Peli sur une image couleur, sur des vidéos et sur des HMD, en demandant aux patients de détecter un événement particulier dans une scène. Cette équipe a utilisé différentes techniques dont la détection d'un événement imprévu, la mesure de l'aisance pour réaliser une tâche, telles que la lecture [57], la déambulation le long d'un couloir [28], l'analyse des performances pour la reconnaissance d'obstacle [55].

Afin d'analyser l'impact des caractéristiques importantes pour l'usage convenable d'une aide par les patients, différentes équipes de recherche ont mesuré le taux d'avis favorables et défavorables pour chaque critère possible de choix [146], et évalué l'apport de chaque

équipement et leurs fonctions utiles [149, 28, 187]. Il peut également être demandé à la personne de réaliser un réglage par elle-même et d'enregistrer cette valeur [179].

Peli a montré que l'affichage de contours ne perturbe pas la détection d'évènements imprévus dans l'environnement [10, 9].

La présentation d'images ou de séquences vidéo permet d'obtenir l'avis d'un certain nombre de patients. Par exemple, il a été présenté une vidéo en symétrie avec une partie traitée et une partie non traitée, le patient devait indiquer la partie la plus pertinente en fonction de ses besoins visuels. Afin de ne pas subir de biais lié à la présentation d'un grand nombre d'images et la mémorisation des images par le patient, une image neutre (image grise uniforme) sépare les séquences ou images.

4.2. Présentation du protocole expérimental

L'objectif de cette évaluation est de répondre à la question suivante :

Le rehaussement des contours peut-il aider les déficients visuels dans l'interprétation d'une scène ?

Après avoir analysé les besoins des déficients visuels, nous avons développé un outil dans le but de répondre à un de ces besoins. Nous allons maintenant évaluer cet outil. Pour cela au vu de l'état de l'art, il semble intéressant de réaliser notre évaluation en demandant aux patients d'indiquer les images les plus faciles à comprendre, les plus adaptées en fonction de leurs pathologies. Ces images sont affichées deux par deux, avec chacune un traitement différent. Dans un but de comparaison, les deux images sont identiques excepté les traitements appliqués à l'une d'elles. Afin que le patient ne soit pas trop influencé par les images précédentes, une image grise neutre est utilisée et les images sont affichées dans un ordre aléatoire.

Notre protocole d'essai doit nous permettre de répondre aux points suivants :

- la ou les méthodes d'extraction de contours à utiliser.
- l'apport de notre pyramide, développée au chapitre précédent,
- l'apport des contours par rapport aux images rehaussées,
- le profil de trait à utiliser pour afficher les contours
- la couleur du trait à utiliser ou la palette de couleurs à utiliser

- l'épaisseur du trait à utiliser

4.2.1. Evaluation des contours

Nous avons mis en place un essai au sein de l'ARAMAV⁵¹, un centre de réadaptation de malvoyants en France et également en collaboration avec Vision et Recherche une entité de Rétina France⁵², une association de malvoyants.

J'ai donc constitué deux groupes de malvoyants : le premier comprenait des patients de l'ARAMAV, le deuxième des adhérents de Rétina France. Ces deux groupes m'ont permis d'utiliser des conditions et matériels de test différents. Cela m'a permis d'analyser l'impact du niveau lumineux sur les réponses, ainsi que l'impact du type d'écran utilisé.

Cette première partie du protocole expérimental avait pour objectif d'évaluer l'apport du rehaussement des contours en fonction des méthodes d'extraction (Sobel, Canny, Pyramide, et ColorCanny). J'ai présenté aux malvoyants un couple d'images. Chaque image était affichée en plein écran. Le sujet pouvait afficher alternativement chacune des images de ce couple autant de fois qu'il le souhaitait. Une fois la réponse du sujet enregistrée le test passait automatiquement au couple d'images suivant. Le sujet disposait d'un temps maximal de 2 minutes par couple d'images pour répondre. Ce temps est suffisamment long pour ne pas stresser le sujet et permet de rester dans une durée maximale raisonnable pour le test complet.

Ce protocole expérimental a été réalisé dans :

- Une pièce :
 - Groupe 1 : une pièce noire. La pièce noire permet de s'affranchir des perturbations extérieures et du niveau de lumière à contrôler pendant les tests. Le malvoyant avait besoin d'un temps d'adaptation avant de débiter le test. Ce temps a été utilisé pour expliquer le fonctionnement du test sur les 6 couples d'images prévus à cet effet.
 - Groupe2 : une pièce éclairée. En journée, cette pièce était éclairée par de la lumière naturelle, store levé et volet ouvert, lorsque la luminosité extérieure baissait les lumières étaient allumées. Le niveau lumineux de la pièce restait supérieur à 200 cd/m2, la recommandation pour un travail sur écran⁵³.

⁵¹ <http://www.aramav.com/pilot.html> clinique présentée au chapitre 2

⁵² <http://www.retina.fr/>

⁵³ <http://www.inrs.fr/accueil/demarche/savoir-faire/eclairage.html>

- Une manette de jeu (reliée au PC de test), cf. Figure 4.1, permet au malvoyant de piloter la présentation des images et de répondre aux questions posées. Cette manette permet de saisir de manière automatique les résultats. Les données collectées sont le choix du malvoyant, le nombre de fois où il a visionné chaque image, le temps de réponse.
- Un écran : la présentation des images était réalisée sur un écran couleur pour tous les patients. Cet écran était placé de manière à représenter idéalement un champ de 40 degrés horizontalement et de 30 degrés verticalement. Le choix du champ de 40 x 30° correspond au plus grand champ d'affichage disponible sur les produits Laster standard. Il était nécessaire que l'écran soit visible de manière confortable par chaque patient.
 - Groupe 1 : un écran CRT couleur, cf. Figure 4.1, était utilisé car il présente l'avantage d'offrir une restitution des couleurs même lorsque le regard est de biais par rapport à sa surface.
 - Groupe 2 : un écran de type LCD était utilisé.
- La position du patient. Le patient était installé confortablement afin de bien voir l'écran, la hauteur des lignes de regard devait correspondre à la hauteur du milieu de l'écran. De plus cette assise permettait au patient de rester attentif pendant toute la durée du test (jusqu'à 30 voire 45 minutes). Le patient se plaçait de manière à voir confortablement l'écran. La distance entre le patient et l'écran était relevée par l'expérimentateur, et devait être proche de celle permet d'obtenir 40° par 30°.
- Lot d'images : chaque lot d'images était constitué de 6 images sélectionnées manuellement, cf. Figure 4.2, par nos soins parmi notre base d'images utilisée au chapitre précédent. Les images qui la composent ont été acquises, avec les systèmes d'acquisitions actuellement possédés par la société Laster Technologies (caméra misumi 3T et caméra USB-Omnivision), en résolution VGA (640 X 480 pixels) et en couleurs. Les images ont été sélectionnées de manière à placer dans chaque lot une image d'extérieur, de rue, d'intérieur, avec du texte et de nuit de manière à obtenir des lots identiques en terme de situations.

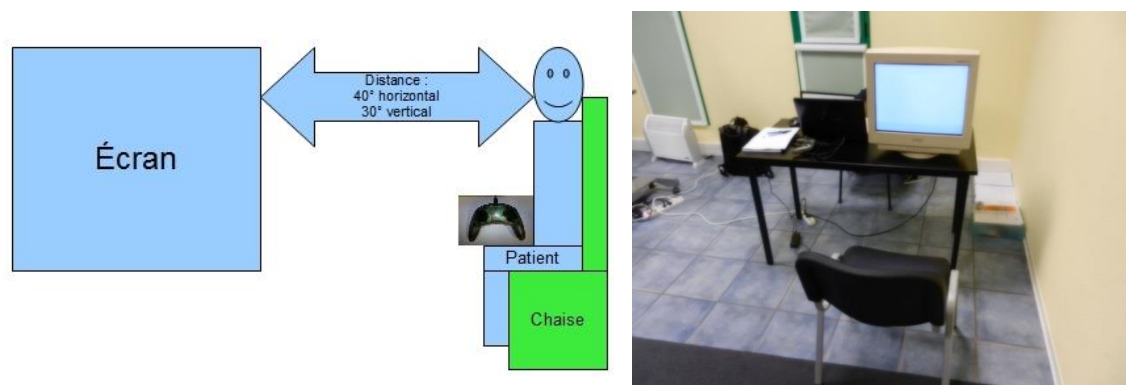


Figure 4.1 : Schéma de côté du dispositif expérimental (à gauche) une photographie du dispositif (à droite)

Pour cette présentation d'images, les algorithmes proposés étaient Sobel, Canny notre pyramide, Canny Couleur et le rehaussement, ces outils sont présentés au chapitre 3 relatif aux développements algorithmiques. Le rehaussement⁵⁴ utilisé est un étirement de la dynamique de l'image, linéaire, afin de rendre plus sombre les éléments sombres et plus lumineux les éléments lumineux, dans le but d'augmenter les différences entre les éléments sombres et clairs de l'image.

Le nombre de couples d'images a été réparti comme suit, en fonction des traitements appliqués :

- Image source, image rehaussée par le filtre de Sobel
- Image source, image rehaussée par le détecteur de Canny
- Image source, image rehaussée par la méthode pyramidale
- Image source, image rehaussée par le Canny Couleur
- Image rehaussée par le filtre de Sobel, image rehaussée par le détecteur de Canny
- Image rehaussée par le filtre de Sobel, image rehaussée par la méthode pyramidale
- Image rehaussée par le filtre de Sobel, image rehaussée par le Canny Couleur
- Image rehaussée par le détecteur de Canny, image rehaussée par la méthode pyramidale
- Image rehaussée par le détecteur de Canny, image rehaussée par le Canny Couleur
- Image rehaussée par la méthode pyramidale, image rehaussée par le Canny Couleur

⁵⁴ <http://www.mathworks.fr/products/matlab/>

- Images identiques rehaussées avec soit le filtre de Sobel, le détecteur de Canny, la méthode pyramidale, le Canny Couleur ou sans rehaussement
- Image source, image rehaussée par étirement de l'histogramme

Il faut donc 11 couples d'images et un couple d'images « image source-image rehaussée » soit 12 couples d'images au total. J'ai utilisé 4 acuités visuelles (10/10, 1/10, 1/20 et 1/50) prédéfinies pour déterminer l'épaisseur des traits pour le rehaussement des contours. Chaque lot devait comporter 6 images différentes, j'ai utilisé deux fois chaque image. Les 12 couples d'images pouvaient être présentés 4 fois à la personne avec 4 acuités différentes.

Pour ce test, les contours détectés étaient superposés à la scène initiale. La couleur jaune était attribuée aux contours. Un tel choix est induit par le fait que le jaune n'est pas filtré par la macula et que ce rayonnement excite les cônes M et L très nombreux dans la rétine humaine.

Il était énoncé au patient :

"Deux images de la même scène vont vous être présentées l'une après l'autre. Vous devrez préciser quelle image de ce couple vous permet de visualiser le plus facilement la scène photographiée. Vous pouvez afficher les deux images alternativement autant de fois que vous voulez. Si vous jugez que les deux images se valent, vous indiquez le choix « pas de différence ». Vous appuyez sur « L » pour afficher la première image, « R » pour afficher la deuxième image. Une fois votre choix fait, vous nous indiquez : l'image A, l'image B ou les deux sont identiques. La première image affichée sera l'image A."

Les malvoyants ont tous visualisé 6 couples d'images pour appréhender le fonctionnement du test avant de réaliser les 48 couples d'images du test effectif. Entre chaque couple d'images était affiché un écran gris neutre pendant 5 secondes.

4.2.2. Evaluation de la restitution

La deuxième partie du protocole d'évaluation devait mettre en évidence les paramètres utiles pour l'affichage des contours. J'ai présenté une image au malvoyant qui devait régler l'épaisseur de trait, la couleur idéale selon lui, ainsi que le mode de représentation des contours.

Cette partie se déroulait de la même manière avec le même matériel, les mêmes conditions et les mêmes images. Le lot d'images pouvait changer entre le test « Evaluation des contours » et le test « Evaluation de la restitution », il était choisi de manière aléatoire pour chaque test. J'ai énoncé au patient :

"Une image contenant des informations ajoutées est affichée devant vous. Vous allez utiliser la manette mise à votre disposition pour adapter à votre convenance les informations ajoutées. Vous pouvez afficher les deux images alternativement autant de fois que vous voulez. Si vous jugez que les deux images se valent, vous indiquez le choix « pas de différence ». Vous appuyez sur « L » pour afficher la première image, « R » pour afficher la deuxième image. Une fois votre choix fait, vous nous indiquez l'image A, l'image B ou les deux sont identiques. La première image affichée sera l'image A. "

Ici, le patient devait commencer par choisir l'épaisseur des contours, puis indiquer parmi les couleurs proposées celle qui lui convenait. Pour cela, les images étaient présentées en couples deux par deux. Enfin le patient devait avec le dernier couple d'images indiquer le profil de contours qui lui convenait le mieux. J'ai proposé aux malvoyants deux types de traits appelés profil, soit un profil en créneau (bande de même intensité sur toute la largeur), soit un profil gaussien (bande plus lumineuse au centre et atténuée sur les bords en suivant la forme d'une gaussienne). La personne devait indiquer toutes les réponses pour une même image puis le test passait à l'image suivante.

4.3. Malvoyants ayant participé à l'évaluation

Les malvoyants du groupe 1, sont au nombre de 28, il s'agit des patients de l'ARAMAV. Les malvoyants du groupe 2 sont des adhérents de Rétina France, au nombre de 30.

Par la suite, le groupe 1 désigne les malvoyants interrogés dans le cadre de l'ARAMAV. De même, le groupe 2 désigne les malvoyants interrogés dans le cadre de Rétina France. Lorsque les données issues des deux groupes de malvoyants ont été fusionnées, il est indiqué « global ».

J'ai analysé chaque paramètre caractéristique de chaque malvoyant et chaque réponse. Lorsqu'on analyse l'âge des malvoyants, cf. Figure 4.2, il n'y a pas parmi les deux groupes de différences statistiques, j'ai donc considéré les deux groupes comme provenant de la même population de malvoyants.

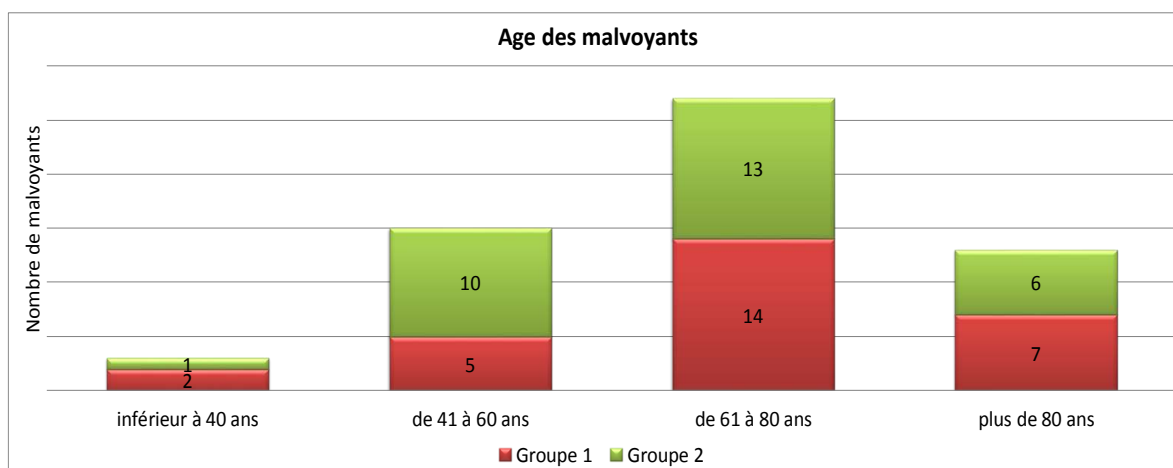


Figure 4.2 Répartition de l'âge des malvoyants ayant répondu au questionnaire

Lorsque l'on compare la répartition des malvoyants en fonction de leur âge entre les deux groupes, le test du khi deux permet de considérer comme un seul et même groupe ces deux populations statistiques, la probabilité pour que l'hypothèse de deux groupes différents est bien inférieure à 0,05.

Lorsque l'on regarde les professions, cf. Figure 4.3, des malvoyants interrogés, il n'y a pas non plus de différences. Il y a bien quelques petites différences, par exemple entre le nombre d'étudiants et le nombre de travailleurs à temps partiels, mais ces différences sont mineures. Le premier groupe était composé de 28 malvoyants, le deuxième groupe était composé de 30 malvoyants. Nous avons donc interrogé 58 malvoyants au total.

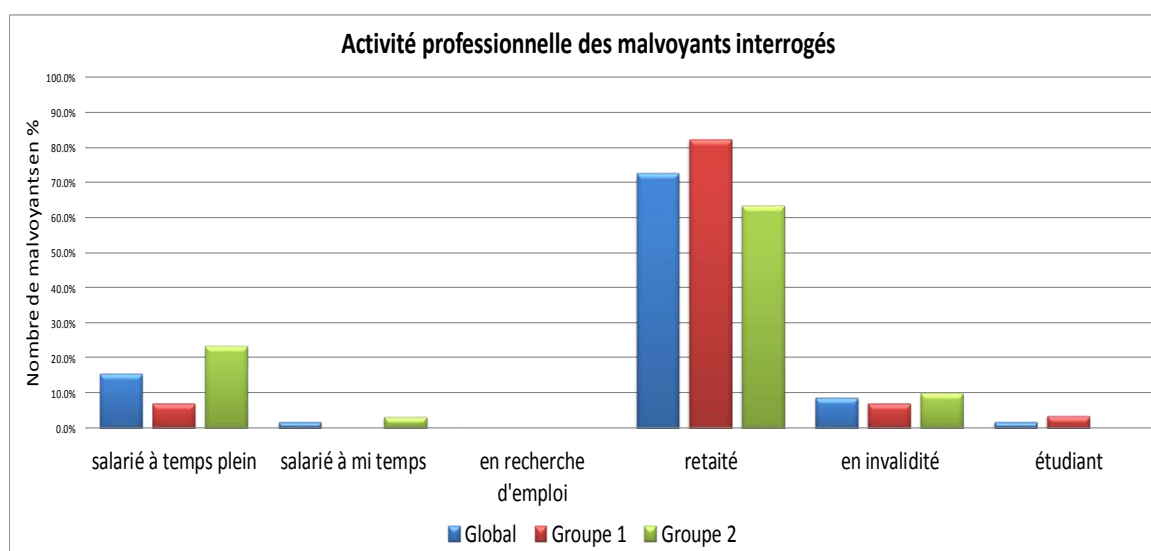


Figure 4.3 Catégorie socio-professionnelles des malvoyants en pourcentage par rapport à chaque groupe.

Lorsque l'on compare la répartition des malvoyants en fonction de leurs catégorie socio-professionnelle entre les deux groupes, le test du khi deux permet de considérer comme un seul et même groupe ces deux populations statistiques, la probabilité pour que l'hypothèse de deux groupes différents est bien inférieure à 0,05.

Certains malvoyants ont expliqué avoir pris leur retraite lors de l'apparition de leur maladie ou avoir été placé en invalidité lors de l'apparition de celle-ci.

Une des informations les plus importantes lorsque l'on parle de malvoyance est l'acuité visuelle. Nous pouvons voir sur la Figure 4.4 la répartition des acuités en vision de loin, exprimées en dixièmes, de chaque œil pour chacun des malvoyants. Pour un certain nombre de malvoyants, l'acuité à l'œil gauche est différente de celle de l'œil droit. De nouveau, au vu de ce critère, les deux groupes sont sensiblement identiques.

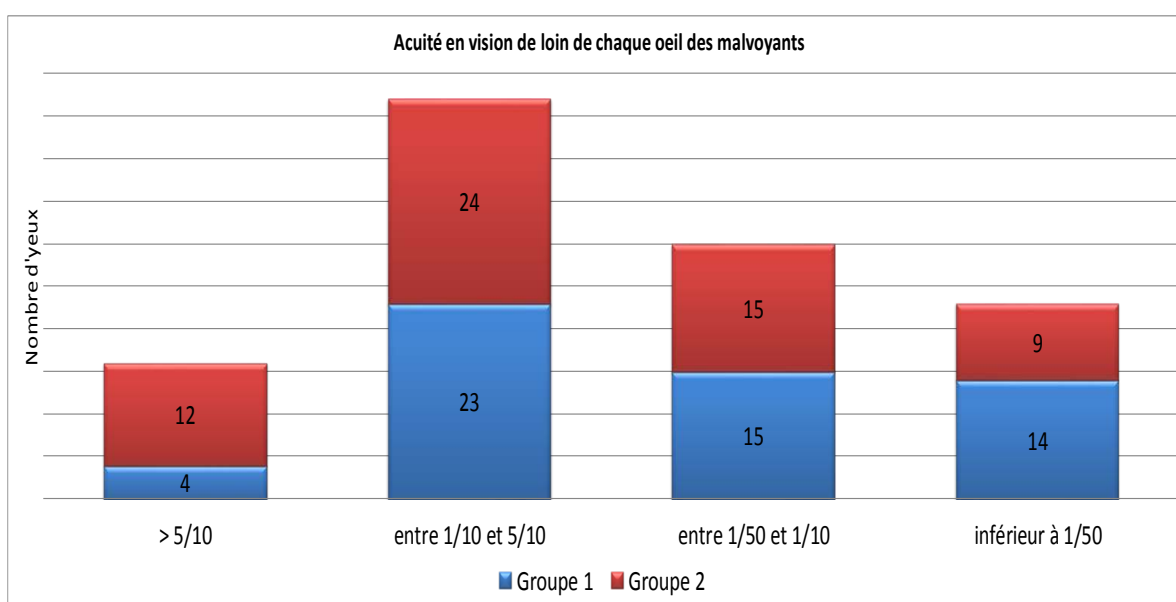


Figure 4.4 Répartition des acuités monoculaires en vision de loin.

Lorsque l'on compare la répartition des malvoyants en fonction de leurs acuités monoculaires entre les deux groupes, le test du khi deux permet de même de considérer comme un seul et même groupe ces deux populations statistiques.

Le deuxième paramètre important est le type d'atteinte du champ visuel. Nous avons pour cela classé les malvoyants en fonction de leur atteinte : scotome, tubulaire, excentré ou autre. Les atteintes de type « excentrement » et « autre » peuvent être cumulées avec une autre atteinte sur la Figure 4.5.

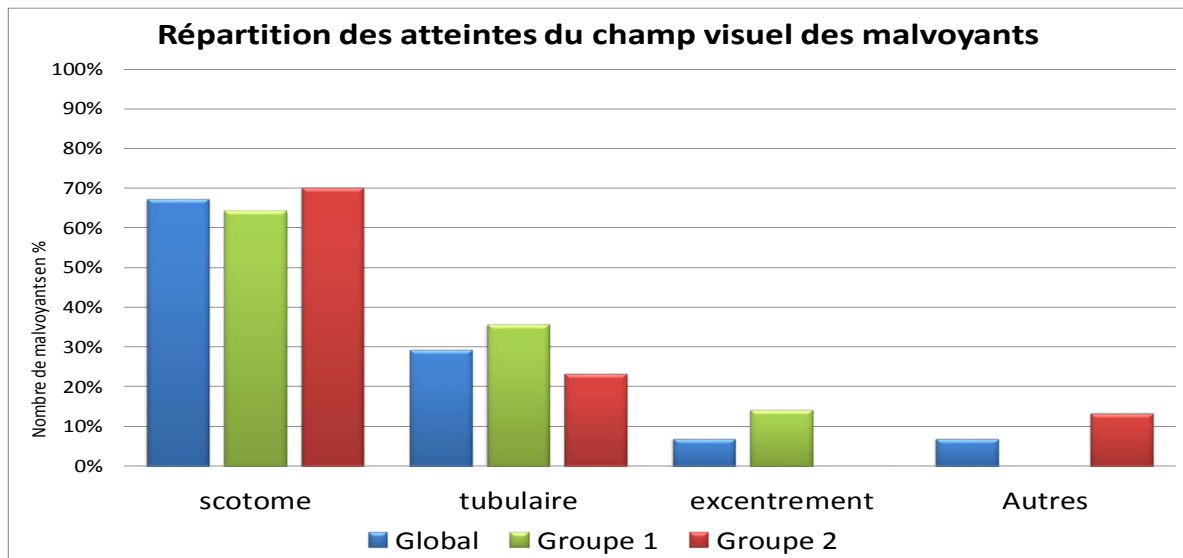


Figure 4.5 Répartition des malvoyants en fonction de l'atteinte du champ visuel.

Lorsque l'on compare la répartition des malvoyants en fonction de leur atteinte du champ visuel entre les deux groupes, le test du khi deux permet de même de considérer comme un seul et même groupe ces deux populations statistiques.

Pour rappel, le scotome concerne la DMLA, il s'agit de la destruction de la partie centrale du champ visuel correspondant à la fovéa et parfois à la macula. L'atteinte tubulaire, concerne la rétinite pigmentaire, il s'agit d'un rétrécissement du champ visuel de l'extérieur vers l'intérieur de celui-ci. Dans ce cas, le malvoyant voit comme à travers un trou de serrure. L'excentrement correspond au cas où la vision tubulaire ou le scotome ne sont pas axés sur le centre du champ visuel, mais décalés par rapport à celui-ci. Le cas de l'atteinte autre correspond à des pathologies n'affectant pas le champ visuel tel que l'albinisme.

4.4. Résultats de l'évaluation

4.4.1. Introduction

Le premier groupe a réalisé le test dans les conditions dits « groupe 1 » et sera appelé groupe 1 pour la suite de ce chapitre. De même le second groupe a réalisé le test suivant les conditions dites « groupe 2 » et sera appelé groupe 2.

Le groupe 1 comprenait 28 malvoyants, mais seulement 25 ont réussi à réaliser le test. Le groupe 2 comprenait 30 malvoyants, mais seulement 28 ont réussi à réaliser le test. J'ai seulement pris en compte pour l'analyse, les malvoyants ayant réalisé le test dans son

ensemble. Certains malvoyants n'ont pas réussi à réaliser ce test intégralement car leurs acuités visuelles étaient trop faibles.

4.4.2. Le rehaussement des contours peut-il être une méthode d'assistance visuelle ?

Nous allons, dans un premier temps, analyser le nombre de malvoyants ayant exprimé au moins une fois une réponse positive avec un rehaussement de contours. Pour au moins un couple d'images Source-Contours, sans différencier la méthode de détection de contours, le malvoyant avait indiqué une préférence pour l'image contours. Lorsqu'une seule taille de rehaussement de contours et un seul détecteur de contours convient au malvoyant, celui-ci peut exprimer un avis favorable pour deux couples de traitements uniquement pendant ce test. De plus, son besoin peut varier d'une image à l'autre pendant ce test. Il est tout à fait possible de ce fait que pour les deux couples de traitements, lui correspondant, que pour l'une des images, il n'est pas besoin du rehaussement. Nous avons donc analysé le nombre de malvoyants ayant exprimé au moins une préférence pour les contours. Nous obtenons un taux de 75,7 % des malvoyants indiquant au moins une préférence pour le rehaussement des contours, dont :

- Pour le groupe 1, un taux de 80 % soit 20 malvoyants,
- Pour le groupe 2, un taux de 71,4 % soit 20 malvoyants.

Dans le groupe 2, il y avait 4 malvoyants possédant encore des acuités visuelles supérieures à $6/10^{\text{ème}}$, l'aide apportée par le rehaussement des contours était faible pour eux. Lorsque l'on analyse ces résultats pour uniquement les individus ayant des acuités adaptées ce test, nous obtenons un taux global de 81,6 % des malvoyants indiquant au moins une préférence pour le rehaussement des contours, dont :

- Pour le groupe 1, un taux de 80 % soit 20 malvoyants,
- Pour le groupe 2, un taux de 83,33 % soit 20 malvoyants.

Ces 4 malvoyants présentaient des acuités supérieures à $6/10^{\text{ème}}$ à au moins un de leurs yeux, Ces 4 malvoyants seront pris en compte pour la suite des données présentées.

Le rehaussement de contours apporte une assistance visuelle aux malvoyants ayant une acuité visuelle inférieure à $6/10^{\text{ème}}$ et supérieure ou égale à $1/50^{\text{ème}}$. Au-dessous de cette acuité

visuelle, la personne aura beaucoup de difficultés à voir les contours et le rehaussement de ceux-ci.

4.4.3. Existe-t-il un lien entre l'épaisseur des traits à utiliser pour le rehaussement des contours et l'acuité visuelle ?

Afin d'analyser s'il existe un lien entre l'acuité et l'épaisseur du trait pour le rehaussement, nous allons dans un premier temps analyser les réponses du groupe 1 puis dans un deuxième temps celles du groupe 2. Nous allons analyser le taux de réponses favorables en fonction de l'épaisseur du rehaussement, pour représenter les contours, lors de l'évaluation. Afin de générer les images 4 épaisseurs de rehaussement, représentant 4 acuités différentes, ont été utilisées. La Figure 4.6 représente les taux de réponses favorables pour les couples d'images utilisant uniquement l'épaisseur de rehaussement, la plus proche des plus petits détails discernables par le malvoyant en fonction de son acuité visuelle. La Figure 4.7 représente les taux de réponses favorables pour les couples d'images utilisant uniquement l'épaisseur de rehaussement correspondant à la taille des plus petits détails discernables par le malvoyant s'il avait une acuité une fois plus élevée, parmi les 4 acuités utilisées pour l'évaluation (trait plus fin). La Figure 4.8 représente les taux de réponses favorables pour les couples d'images utilisant uniquement l'épaisseur de rehaussement correspondant à la taille des plus petits détails discernables par le malvoyant s'il avait une acuité une fois plus faible, parmi les 4 acuités utilisées pour l'évaluation (trait plus gros). Les graphiques suivants représentent les taux de réponses favorables, neutres ou défavorables pour chaque traitement. Les Figures 4.6, 4.7 et 4.8 représentent le taux de réponses favorables (positif), défavorables (négatif) et neutres (égal) pour chaque couple de traitement.

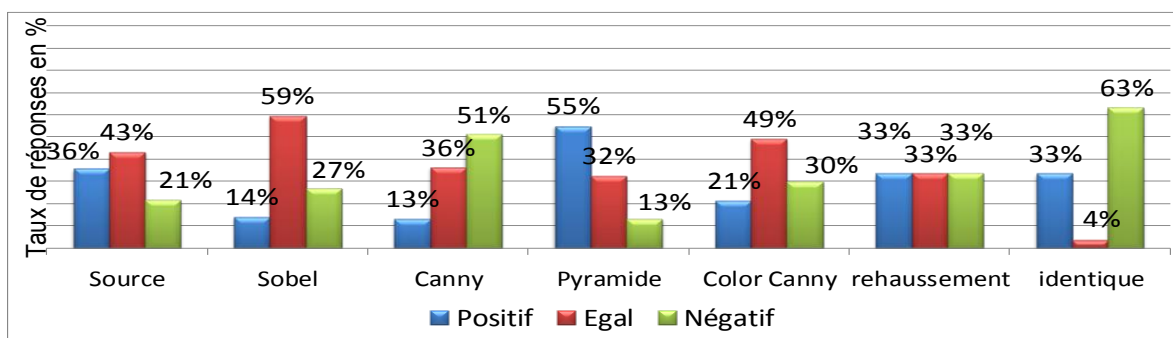


Figure 4.6 : Taux de réponses pour chaque méthode uniquement pour l'épaisseur de rehaussement le plus proche de l'acuité de chaque malvoyant

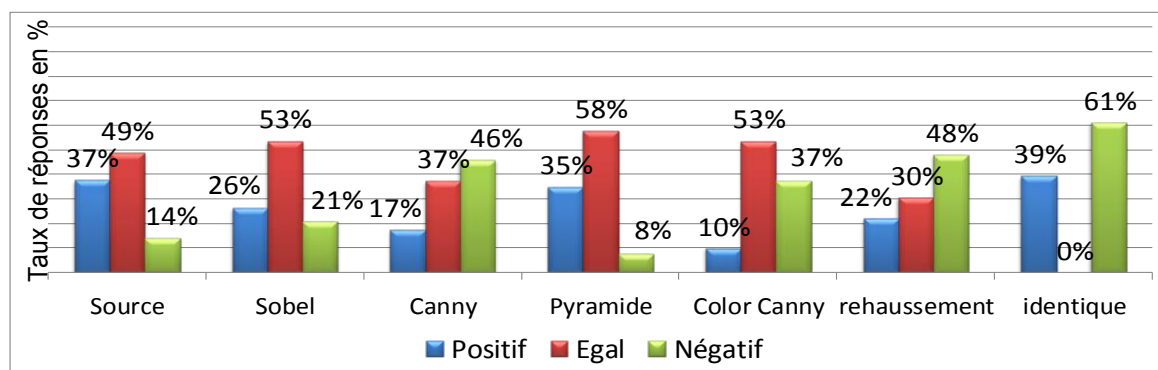


Figure 4.7 : Taux de réponses pour chaque méthode uniquement l'épaisseur de rehaussement le plus proche de l'acuité une fois supérieure de chaque malvoyant

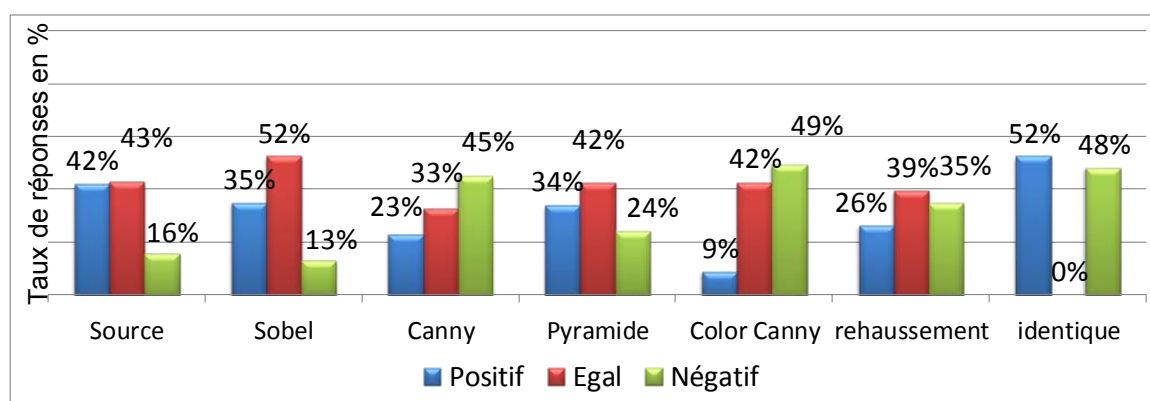


Figure 4.8 : Taux de réponses pour chaque méthode uniquement pour l'épaisseur de rehaussement le plus proche de l'acuité une fois inférieure de chaque malvoyant

Il apparaît entre les Figures 4.6 et 4.7 et entre les Figures 4.6 et 4.8 des différences entre les taux de réponses favorables pour chaque traitement. Les malvoyants pour certaines images vont préférer des traits à peine discernables mais suffisants pour créer une modification de leur perception, dans d'autres cas, ils vont préférer des traits beaucoup plus marqués. L'épaisseur des traits doit être adaptée en fonction de l'acuité.

Les taux de réponses favorables de la Figure 4.6 montrent que le rehaussement de contours est plus performant pour les malvoyants lorsque les traits de restitution sont adaptés à l'acuité de la personne. Il apparaît également une variation du taux de réponses favorables en fonction de l'épaisseur du trait de restitution utilisée. Il existe bien un lien entre l'acuité et les performances du rehaussement de contours, ce lien n'est pas simple, il faut prendre en compte l'image et l'épaisseur de trait de restitution.

Les Figures 4.9, 4.10 et 4.11 représentent respectivement mêmes informations, les Figures 4.6, 4.7 et 4.8, pour le deuxième groupe de malvoyants.

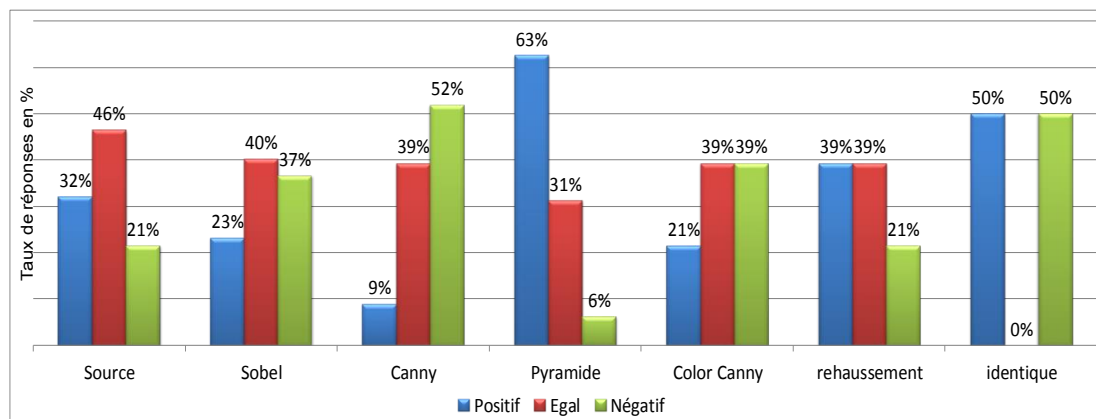


Figure 4.9 : Taux de réponses pour chaque méthode uniquement pour l'épaisseur de rehaussement le plus proche de l'acuité de chaque malvoyant

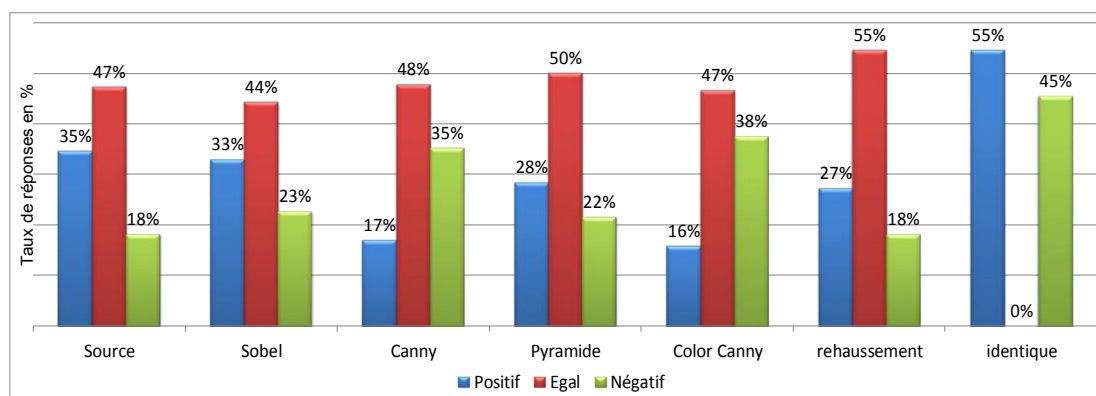


Figure 4.10 : Taux de réponses pour chaque méthode uniquement l'épaisseur de rehaussement le plus proche de l'acuité une fois supérieure de chaque malvoyant

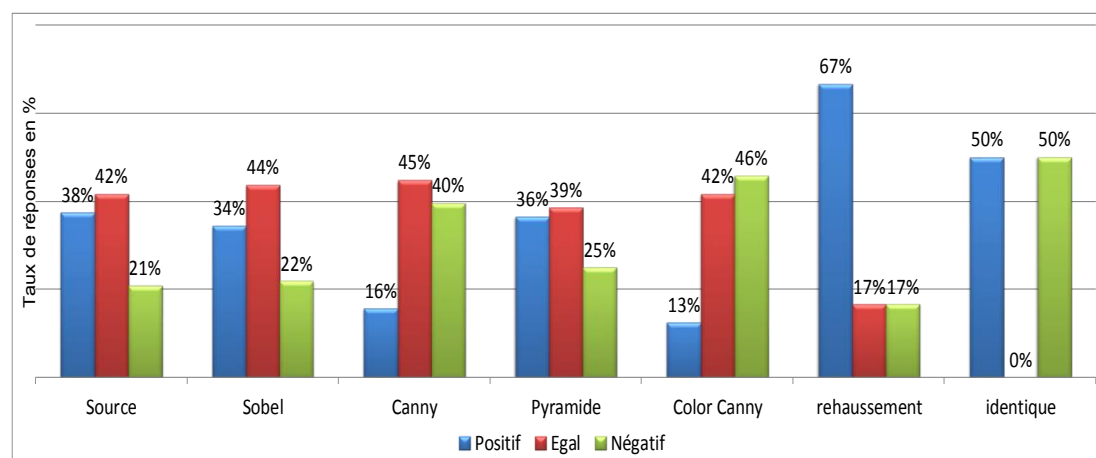


Figure 4.11 : Taux de réponses pour chaque méthode uniquement pour l'épaisseur de rehaussement le plus proche de l'acuité une fois inférieure de chaque malvoyant

Pour les deux groupes, nous obtenons les mêmes conclusions concernant le lien entre l'acuité visuelle et les performances du rehaussement de contours.

4.4.4. Quelles sont les méthodes d'extraction de contours les plus pertinentes pour l'assistance des malvoyants ?

Afin de répondre à cette question, nous avons analysé le taux de réponses favorables au rehaussement de contours. Tous les malvoyants ont été pris en compte pour la suite des analyses. Autrement dit, nous avons analysé le nombre d'images pour lesquelles chaque malvoyant exprimait une préférence pour le rehaussement de contours par rapport aux images sans rehaussement. Nous avons, sur la Figure 4.12, le taux pour chaque extracteur de contours. Comme indiqué dans le protocole, section 4.2, les images ont été générées pour 4 épaisseurs de traits différentes. Les « données brutes » représentent le taux de réponses quel que soit l'épaisseur du trait de rehaussement. Les données « meilleure acuité » indiquent ces mêmes informations mais uniquement pour les images dont l'épaisseur de rehaussement était le plus proche de l'acuité du malvoyant, dans ce cas, l'épaisseur cette épaisseur est proche de la taille du plus petit élément discernable par le malvoyant.

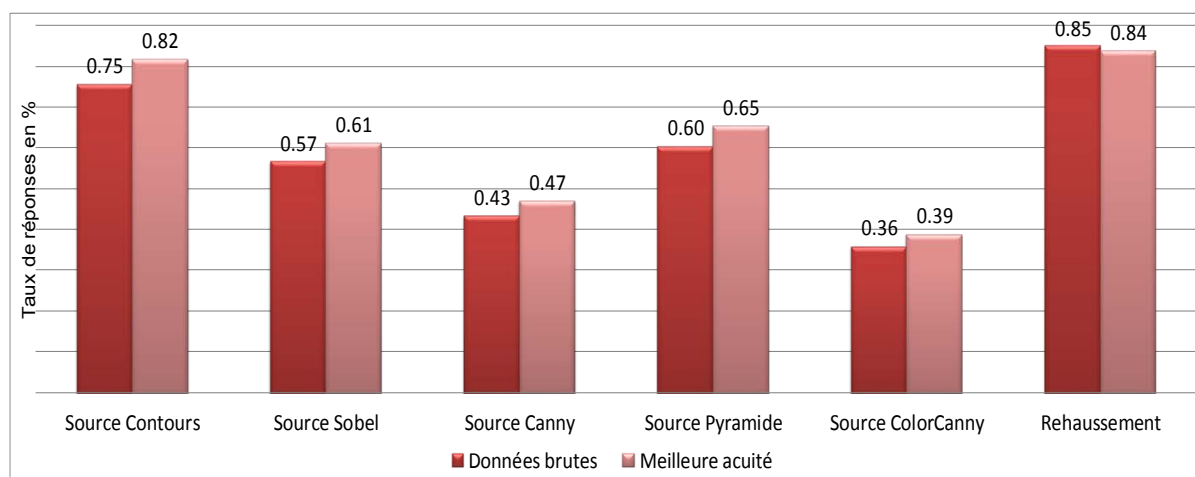


Figure 4.12 : Taux de réponses favorables au rehaussement des contours par rapport aux images sans traitement

La Figure 4.12, représente le taux de préférence de chaque détecteur de contours par rapport aux images sans rehaussement de contours. Les images de chaque couple de détecteurs de

contours ont été présentées alternativement dans l'ordre et en sens inverse, afin de ne pas avoir de biais dû à l'ordre des images. Comme nous pouvons le voir sur la Figure 4.12, nous retrouvons le taux de 82 % de réponses favorables pour le rehaussement de contours. Il est possible de réaliser un classement des extracteurs de contours du plus performant au moins performant comme suit : la Pyramide, Sobel, Canny et ColorCanny. J'ai réalisé des tests statistiques, pour chaque couple de traitement afin d'analyser l'impact sur les résultats :

- Des conditions de test. Lorsque l'on analyse les valeurs toutes les valeurs avec une loi binomiale, celles-ci peuvent être considérées comme une seule et même série pour un taux de confiance de 95 %. On peut en déduire qu'il n'y a pas d'influence des conditions de test. Il n'y a pas d'influence de l'écran ou du niveau lumineux de la pièce.
- Du type d'atteinte du champ visuel. Lorsque l'on analyse les valeurs toutes les valeurs avec une loi binomiale, celles-ci peuvent être considérées comme une seule et même série pour un taux de confiance de 95 %. Il n'y a pas d'influence du type d'atteinte de champ visuel sur l'apport des contours. Pour le rehaussement, lors que l'on analyse les valeurs toutes les valeurs avec une loi binomiale, en revanche, celles-ci ne peuvent pas être considérées comme une seule et même série pour un taux de confiance de 95 %, cela signifie que le type d'atteinte du champ visuel a une influence.

Nous avons analysé pour chaque paramètre, si les données des groupes sont suffisamment proches pour que l'on considère les deux en tant que un seul groupe (aucune influence du paramètre testé). Par exemple, j'ai analysé les données du groupe 1 et 2, ayant des conditions de test différentes, pour chaque couple de méthodes.

Les extracteurs de contours présentent de fortes différences les uns par rapport aux autres. Sur la Figure 4.13, nous avons le taux de réponses favorables pour les extracteurs de contours par rapport aux images source comme sur la Figure 4.12 mais également le taux de réponses favorables pour la deuxième méthode de chaque couple de traitement.

Le détecteur de contours de Canny est faible par rapport à la méthode de Sobel (Sobel-Canny) et la Pyramide (Canny-Pyramide), cf. Figure 4.13. Le détecteur de Canny et le détecteur ColorCanny sont quasiment équivalents en termes d'avis favorables. L'extracteur de contours de Sobel présente des performances intéressantes par rapport aux méthodes de Canny et ColorCanny. La pyramide est quant à elle, la méthode la plus performante par rapport aux méthodes testées ici. Il existe une nette différence entre les taux de réponses favorables pour

la pyramide et les autres méthodes (Sobel-Pyramide, Canny-Pyramide, Pyramide-ColorCanny).

Le rehaussement présente un taux d'avis favorable élevé, il est nécessaire de modifier toute l'image. L'affichage d'une image complète en réalité augmentée peut engendrer des troubles de la mobilité.

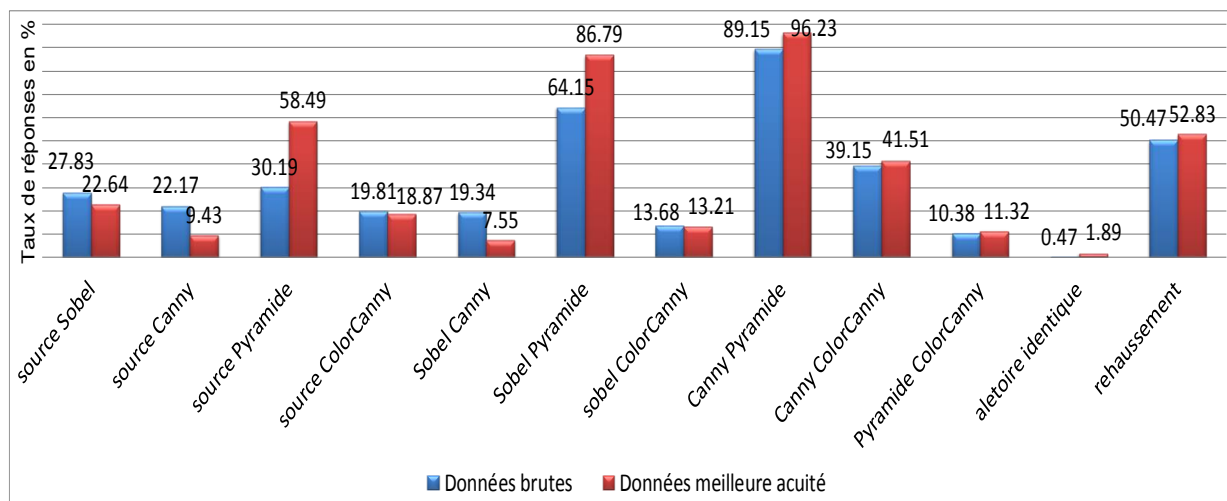


Figure 4.13 : Comparaison du taux de réponses favorables pour chaque méthode par rapport aux images sources, ou de la 2^{ème} méthode du couple par rapport à la 1^{ère} méthode du couple de traitements.

Il est ainsi possible d'établir un classement des méthodes d'extraction de contours en fonction de leur taux respectif de réponses favorables : Pyramide, Sobel, Canny, ColorCanny.

4.4.5. Existe-t-il des couleurs préférentielles pour rehausser les contours ?

Pour rehausser, il est important de choisir une couleur adaptée aux besoins des malvoyants en fonction de leur atteinte, du contexte, comme le montre la Figure 4.14 qui présente le taux de sélection de chaque couleur parmi un couple de couleurs proposées. Les malvoyants, ayant un champ de vision résiduelle tubulaire, privilégient le jaune et le blanc principalement, Figure 4.14. Les malvoyants ayant un champ de vision résiduelle avec un scotome préfèrent le jaune et blanc mais préfèrent pour quelques images du vert, du rouge ou du bleu, leurs réponses sont plus variées.

Il existe donc d'une part des couleurs préférentielles pour l'affichage des contours, et d'autre part ces couleurs varient en fonction du type d'atteinte du champ visuel.

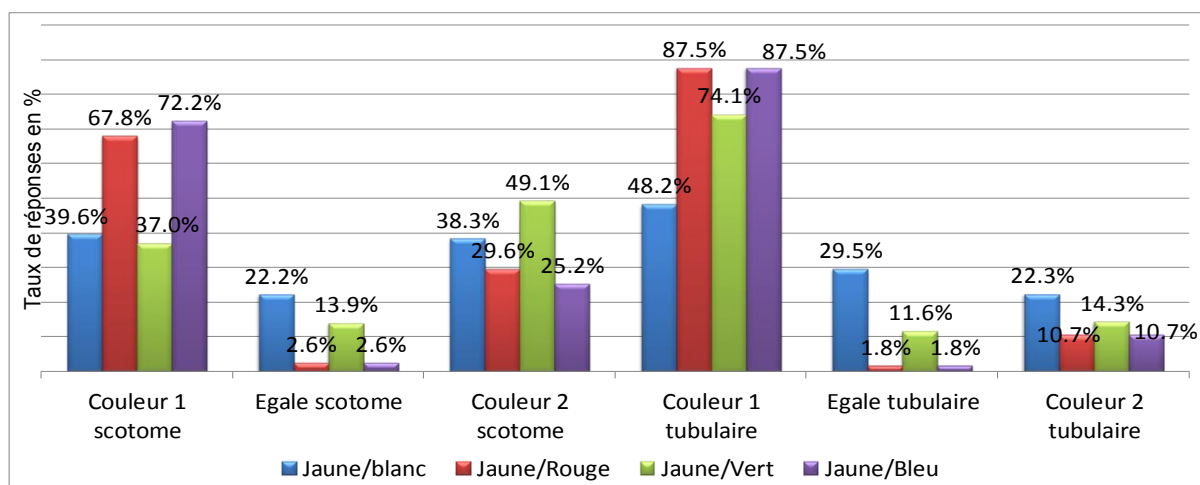


Figure 4.14 : Taux de préférence de chaque couleur de rehaussement en fonction de l'atteinte du champ visuel

Le graphique de la Figure 4.14, présente le taux de réponses favorables à la couleur jaune, ou favorable à la deuxième couleur de chaque couple de couleurs présentées. Les termes en abscisse, « couleur 1 » indiquent une préférence pour la première couleur du couple proposé, par exemple, pour le couple jaune/blanc, le jaune a été préféré.

Lorsque l'on compare les choix de couleurs de rehaussement des malvoyants en fonction des conditions de test entre les deux groupes, le test du khi deux permet de considérer comme un seul et même groupe ces deux populations statistiques, la probabilité pour que l'hypothèse de deux groupes différentes est bien inférieure à 0,01 : les conditions du test n'influencent pas les résultats. Cependant, lorsque l'on compare les choix de couleurs de rehaussement des malvoyants en fonction du type d'atteinte du champ visuel entre les deux groupes, le test du khi deux ne permet pas de considérer comme un seul et même groupe ces deux populations statistiques, la probabilité pour que l'hypothèse de deux groupes différentes est bien supérieure à 0,01 : le type d'atteinte du champ visuel va influencer le choix de la couleur de rehaussement.

Après avoir analysé les préférences pour chaque couleur des malvoyants, il est intéressant d'analyser si ces préférences sont répétables sur une même image. Pour cela nous avons présenté aux malvoyants deux fois chaque couple de couleurs sur une même image. Nous avons analysé le nombre de variations respectives des réponses entre chaque couple d'images. Nous avons présenté deux fois les 4 couples de couleurs sur une même image. Le nombre

maximal d'erreurs peut donc être de 4. Le nombre de variations entre la première et la deuxième analyse des couples couleurs sur une même image est présenté sur la Figure 4.15 pour le groupe 1 et sur la Figure 4.16 pour le groupe 2.

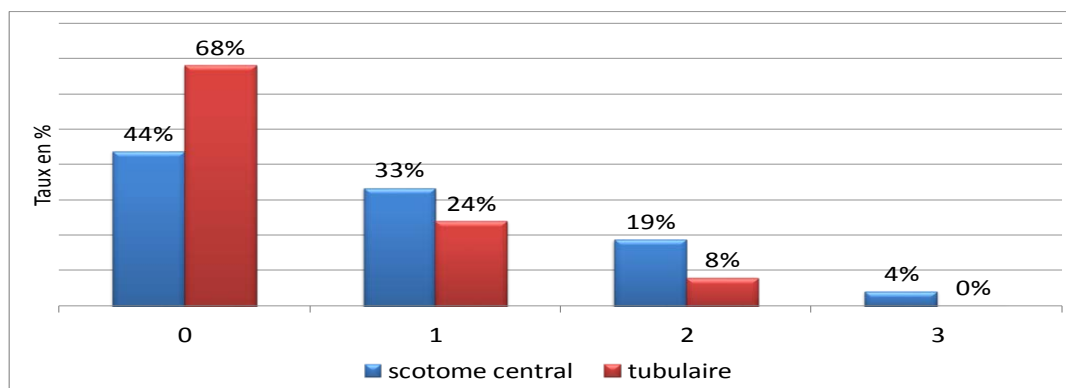


Figure 4.15 : Nombre de divergences de réponses pour un même couple de couleurs pour les malvoyants du groupe 1 en fonction du type d'atteinte du champ visuel

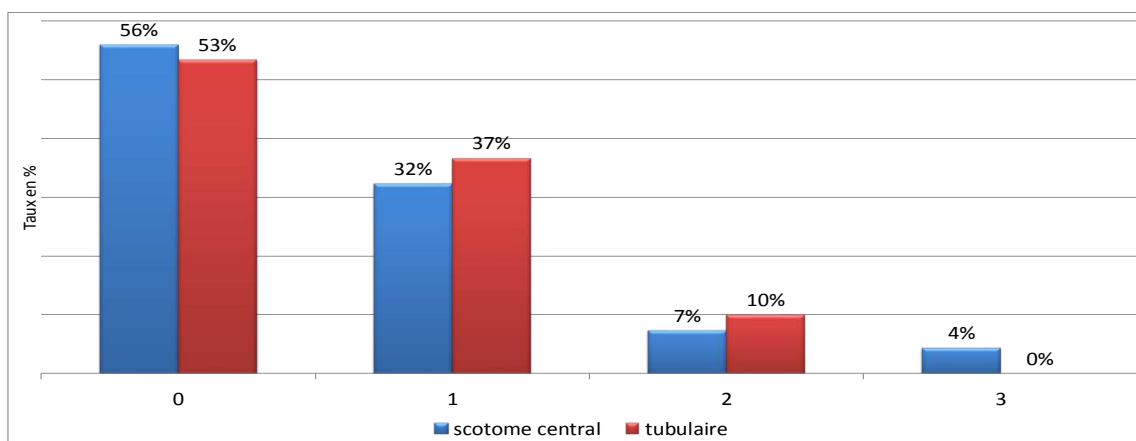


Figure 4.16 : Nombre de divergences de réponses pour un même couple de couleurs pour les malvoyants du groupe 2 en fonction du type d'atteinte du champ visuel

Dans plus de la moitié des réponses quelle que soit la couleur sélectionnée, les malvoyants ont donné les mêmes réponses. On peut en déduire que les malvoyants ont besoin de pouvoir sélectionner la couleur du rehaussement en fonction de la scène et surtout que celle-ci, malgré leur malvoyance, n'est pas sélectionnée de manière arbitraire.

4.4.6. Existe-t-il des profils de traits pour rehausser les contours ?

Après avoir évalué les couleurs pour le rehaussement, j'ai également évalué différents profils d'affichage. Les deux profils de traits utilisés étaient le profil en créneau et le profil gaussien, ils sont présentés à la section 3.2 de ce chapitre.

Les deux profils sont utilisés par les malvoyants lors de cette évaluation, cf. Figure 4.17. Les réponses sont sensiblement identiques quelle que soit l'atteinte du champ visuel des malvoyants. Il a été utilisé deux profils, créneau (bande uniforme) et gaussien (filtre gaussien appliqué à la carte de contours). En abscisse, il est représenté chaque réponse, préférence pour le profil en créneau, préférence pour le profil gaussien ou les deux profils sont égaux pour chaque type d'atteinte du champ visuel.

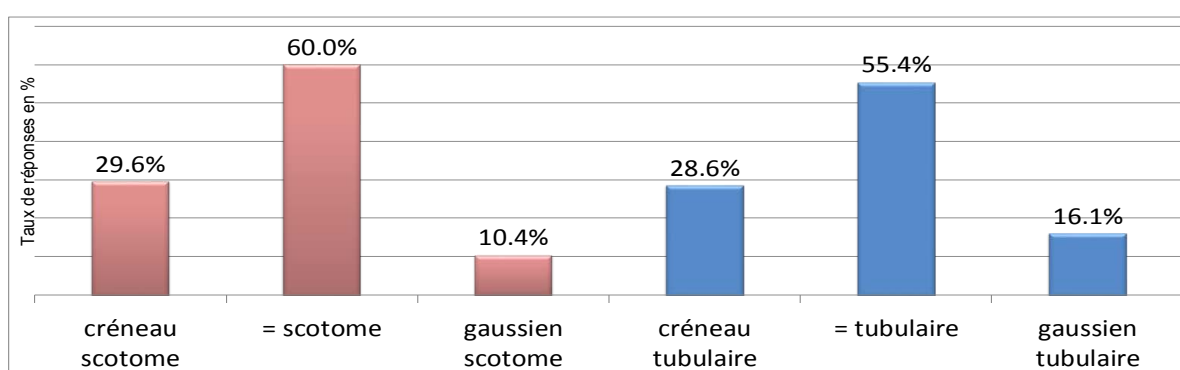


Figure 4.17 : Choix de profil de trait en fonction de l'atteinte du champ visuel

Le graphique de la Figure 4.17, présente le profil de rehaussement choisi par le malvoyant en fonction de l'atteinte du champ visuel, à gauche les réponses des malvoyants dont le champ visuel est atteint de scotome central et à droite les réponses des malvoyants dont le champ visuel est tubulaire.

4.4.7. Existe-t-il des épaisseurs de traits à privilégier pour rehausser les contours ?

Après avoir analysé les paramètres de couleurs et de profil des traits pour rehausser les contours, je vais maintenant analyser l'épaisseur idéale, de ces traits, sélectionnée par les malvoyants. J'ai présenté la valeur moyenne et l'écart type de l'épaisseur des traits en

fonction du type d'atteinte du champ visuel sur la Figure 4.18 pour le groupe 1 et sur la Figure 4.19 pour le groupe 2.

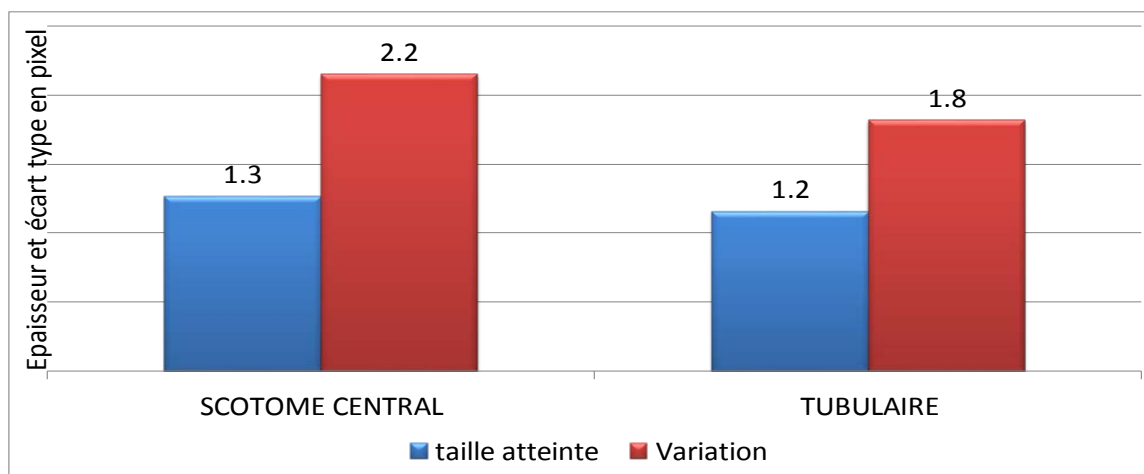


Figure 4.18 : L'épaisseur du trait (en bleu) en fonction de l'atteinte du champ visuel et l'écart type (en rouge) pour le groupe 1.

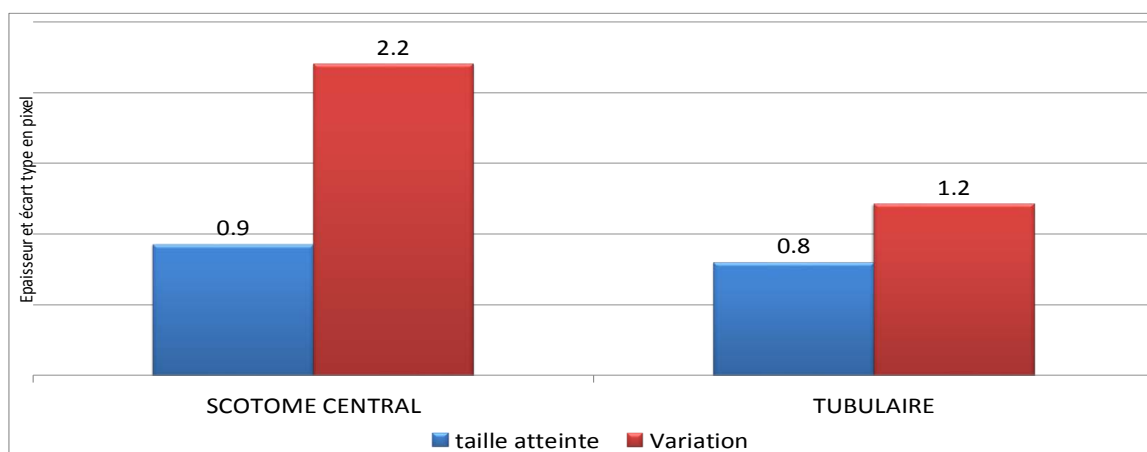


Figure 4.19 : L'épaisseur du trait (en bleu) en fonction de l'atteinte du champ visuel et l'écart type (en rouge) pour le groupe 2.

Pour les deux groupes, l'écart type est plus grand pour les malvoyants ayant un champ visuel comprenant un scotome. Ces malvoyants ont tendance à utiliser des traits plus larges sur certaines images. Les malvoyants atteints de champ de vision tubulaire sollicitent également quelques fois des traits plus larges en fonction de la scène. Enfin, il apparaît une variation de la taille moyenne quelle que soit l'atteinte du champ visuel entre le groupe 1 et le groupe 2, Figure 4.18 et Figure 4.19. Dans des conditions normales d'éclairage (écran d'ordinateur dans une pièce éclairée par la lumière du jour) les malvoyants ont tendance à moins grossir les traits pour la restitution des contours. Ce test ayant été réalisé sur écran, il sera intéressant de vérifier ces informations lors de tests sur le dispositif réel.

4.5. Conclusion de l'évaluation

La réponse à la première question, à savoir si les contours peuvent être utilisés comme méthode d'assistance visuelle pour les malvoyants, est positive (80 % des malvoyants ont au moins une fois exprimé obtenir un apport par les contours). Les contours sont un outil intéressant pour rehausser la réalité des déficients visuels et les aider dans leur interprétation des scènes avec lesquelles ils interagissent.

La réponse à la deuxième question, à savoir quelle est la méthode la plus demandée par les malvoyants parmi les méthodes d'extraction, est la pyramide. Cette réponse valide notre outil par rapport aux études précédentes qui permettaient de montrer l'apport des contours [10, 9, 146, 149, 28, 187] mais avec des méthodes inadaptées aux HMD à verres transparents orientés pour la mobilité [9]. Ces méthodes utilisent aussi des couleurs sombres et très lumineuses pour rehausser les contours ; les couleurs sombres sont peu ou pas restituées sur des dispositifs à verres transparents. Les contours intégrés aux HMD peuvent permettre d'éviter un certain nombre d'obstacles en agrandissant la zone de sécurité de la personne [148, 109].

L'ensemble des paramètres de restitution, couleurs, profil et épaisseur, semble être utilisé par les malvoyants pour afficher les contours sur les images originales.

En effet, il existe d'une part des couleurs préférentielles pour l'affichage des contours, et d'autre part ces couleurs varient en fonction du type d'atteinte du champ visuel. Les malvoyants, atteints de champ de vision tubulaire, privilégient le jaune et le blanc principalement. Les malvoyants atteints de champ de vision avec un scotome préfèrent le jaune et blanc mais peuvent également souhaiter du vert, du rouge ou du bleu, leurs réponses sont plus variées.

Les réponses, pour les deux profils de trait, sont sensiblement identiques quelle que soit l'atteinte du champ visuel des malvoyants. De même, l'épaisseur des traits est adaptée par les malvoyants en fonction des images, de la luminosité, de leur acuité et de leur champ visuel.

Il semble nécessaire de conserver les réglages des paramètres de rehaussement (épaisseur, couleur, profil) lors du développement d'une solution en mobilité et d'une aide pour malvoyant. Ces paramètres pourront être adaptés soit par le malvoyant lui-même soit pré-réglés par un professionnel lors de la configuration initiale du dispositif.

5. Dispositif d'assistance

Afin de réaliser un dispositif d'assistance visuelle adaptatif paramétrable et adapté aux besoins des malvoyants, il est nécessaire de réunir : un dispositif de réalité augmentée pour la restitution, une caméra pour l'acquisition, un système de traitement de l'information et surtout une interface adaptée.

Je vais dans ce chapitre tout d'abord présenter différentes solutions d'affichage afin de montrer quelles sont les solutions possibles et celle que j'ai choisi de développer. Pour cela, j'ai effectué une synthèse des différentes solutions, et recensé leurs avantages et inconvénients. Dans la section 5.1.4, je justifie mon choix de technologie de restitution. Dans la section 5.1.2.3, je vais également présenter le modèle de génération d'image implémenté afin d'obtenir des images binoculaires. Dans la section 5.1.3, je vais présenter la notion de recalage d'information sur la réalité, comment résoudre cette problématique en fonction des technologies et effectuer un état de l'art des techniques permettant de résoudre ou compenser ces problèmes.

Je vais ensuite présenter, dans la section 5.2, les différentes technologies d'acquisition d'images adaptées à notre problématique. Je vais présenter les principaux critères de sélection d'une caméra pour une application de réalité augmentée. Je vais justifier le choix du capteur sélectionné.

Dans la section 5.3, je vais présenter mes recherches de composants et le développement de l'unité de traitement d'image. Je vais présenter l'architecture logicielle, puis le matériel utilisé et enfin les performances ainsi obtenues.

5.1. Dispositif de restitution visuelle

5.1.1. Réalité virtuelle vs Réalité augmentée

Il existe plusieurs catégories de « réalité améliorée », soit la réalité virtuelle soit la réalité augmentée [115].

- La réalité virtuelle : ne permet pas d'obtenir de sensation de la part de l'environnement extérieur ni du monde réel. L'observateur est plongé dans une enceinte ou utilise un équipement qui ne lui permet pas de percevoir l'environnement réel dans lequel il a l'habitude d'évoluer. La Figure 5.1, présente un produit de réalité virtuelle : le Trivisio ARVision 3D est un dispositif binoculaire stéréoscopique de réalité virtuelle (le système est immersif : opaque) qui fonctionne avec deux caméras afin de restituer une image de la réalité.



Figure 5.1 : Exemple de produit de réalité virtuelle : un Trivisio ARVision 3D⁵⁵

- La réalité augmentée : permet d'afficher un contenu visuel qui s'ajoute à la réalité, que l'on peut percevoir comme on a l'habitude de voir. La Figure 5.2 présente 3 produits de réalité augmentée à verres transparents.



Figure 5.2 : Exemples de produits de réalité augmentée : de gauche à droite ; le masque MG1⁵⁶, le Pro Mobile Display⁵⁷ et le Laster See-Thru⁵⁸ de Laster Technologies

Il existe également quelques systèmes entre ces deux catégories, cf. Figure 5.3. Il s'agit de système non ou très peu immersif permettant de visualiser des images et des vidéos sans toutefois permettre d'enrichir la réalité d'information. Ces deux systèmes proposent des champs d'affichages de 15°, ce qui est relativement faible pour ajouter des informations à la

⁵⁵ <http://www.trivisio.com/trivisio-products/arvision-3d-hmd-7/>

⁵⁶ <http://www.laster.fr/produits/promobiledisplay/>

⁵⁷ <http://www.laster.fr/produits/MG1/>

⁵⁸ <http://www.laster.fr/produits/laster-see-thru/>

réalité. Dans les deux cas, ces systèmes ne présentent donc pas les caractéristiques requises (champs d'affichage et transparence) pour réaliser de l'enrichissement de la réalité. Ces dispositifs sont presque inclassables car ils présentent presque toutes les caractéristiques des systèmes de réalité augmentée sans toutefois permettre l'affichage des enrichissements, issus de l'analyse du flux vidéo de la caméra, comme-ci ceux-ci étaient dans la réalité.



Figure 5.3 : Exemples de produits permettant l'affichage d'information : de gauche à droite ; la Google Glass⁵⁹ et le Vuzix M100⁶⁰

Notre système visuel échange des informations sensorielles avec nos autres sens, l'audition et le toucher, afin d'améliorer notre perception de l'espace et notre équilibre. Cela nous permet de mieux nous mouvoir en toute situation, y compris dans des lieux où les informations visuelles sont faibles comme dans une rue non éclairée. Le fait d'utiliser un système de réalité virtuelle perturbe ce mécanisme, car le système visuel humain ne peut plus collecter d'indices extérieurs. Avec un système à réalité augmentée, il est possible de maintenir ce mécanisme, même quand l'utilisateur est en mouvement.

De plus, il est primordial de prendre en compte, lors de la réalisation d'un équipement d'affichage sur tête dit HMD (Head Mounted Display), la distance de vision (distance de travail) pour définir la position du point de convergence (plan de focalisation). En temps normal, il existe une relation linéaire (fonction affine) entre la convergence (en dioptrie : unité optique permettant de quantifier une vergence optique) et la convergence (en dioptrie prismatique : unité permettant de quantifier la déviation d'un axe optique). Les systèmes de réalité virtuelle ne prennent généralement pas en compte cette relation, car l'image est souvent affichée à l'infini alors que l'objet à observer est à une distance finie.

À l'inverse, les systèmes (afficheurs sur tête) de réalité augmentée ne génèrent pas de trouble d'accommodation, et de convergence car la relation accommodation-convergence est conservée [139].

Le design du système a un impact sur le champ de vision, en effet, sa forme et sa taille ont un lien direct avec la taille du champ visuel disponible pour l'observateur à travers le système. En

⁵⁹ <http://www.google.com/glass/start/>

⁶⁰ http://www.vuzix.com/consumer/products_m100.html

effet, la mécanique du dispositif d'affichage occulte les bords du champ visuel. Il est donc nécessaire de réaliser de grands verres et de prévoir en dehors de la zone d'affichage des parties transparentes. De plus, les zones du champ visuel non obstruées ne doivent pas être limitées à la zone couverte par l'afficheur [195].

5.1.2. Dispositifs à verre transparent

Au vu des besoins des personnes malvoyantes et des résultats présentés aux chapitres 2, 3 et 4, il apparaît que la technologie la plus appropriée pour superposer les contours à la réalité et en temps réel est un dispositif de réalité augmentée. Il existe sur le marché un certain nombre de dispositifs monoculaires et/ou binoculaires de réalité augmentée. Ces solutions sont basées sur des technologies proches, quoique légèrement différentes. Nous allons, ci-après, présenter ces différentes solutions d'affichage sur verres transparents pour un usage monoculaire, puis lorsque l'on associe deux modules optiques pour un usage en mode binoculaire [115].

5.1.2.1. Systèmes monoculaires

Les dispositifs de réalité augmentée à verres transparents sont composés de deux sous-ensembles : une source vidéo et les éléments optiques pour conjuguer l'image de la source vidéo à l'image rétinienne perçue par l'utilisateur.

Je vais présenter ci-dessous un certain nombre de sources vidéo, ainsi que leurs principales caractéristiques cf. Figure 5.4.

	CRT	LCD	LCoS	FLCoS	OLED
Taille de la diagonale en pouce	>0,5	>0,7	>0,21	>0,2	>0,66
Durée de vie (heures)	40 000	20 000 – 40 000	10 000 – 15 000	10 000 – 15 000	<10 000
Luminosité (cd/m ²)	100– > 30000	100 <i>Dépend du rétroéclairage</i>	100 – >14000 <i>Dépend du rétroéclairage</i>	100 – >20000 <i>Dépend du rétroéclairage</i>	100 - 2200

Contraste	300 :1 – 700 :1	150 :1 – 450 :1	>2000 :1	>2000 :1	450 :1– >10000 :1
Type d'émission	balayage	rétroéclairage	Rétroéclairage optique	Rétroéclairage optique	Auto émission
Couleurs	Monochrome ou couleurs	Monochrome ou couleurs	Monochrome ou couleurs	Monochrome ou couleurs	Monochrome ou couleurs
Respect du rendu des couleurs sombres	Oui , aucun balayage	Gris résiduel du au rétroéclairage	Gris résiduel du au rétroéclairage	Gris résiduel du au rétroéclairage	Oui pas d'émission

Figure 5.4 : Sources vidéo possibles pour les dispositifs d'affichage sur tête [158].

Le terme « Respect couleur noire » caractérise la bonne restitution de la couleur noire. Lorsque l'utilisateur souhaite afficher un motif noir : absence d'information pour un système à verres transparents, celui-ci être noir (complètement transparent). Il ne doit pas y avoir de voile gris résiduel dû au rétroéclairage de la source vidéo. L'absence de ce voile est importante pour la restitution des couleurs, cela permet de ne pas perturber la vision de l'environnement dans lequel évolue l'utilisateur.

Le deuxième paramètre important est la luminosité de la source vidéo. En intérieur, il est souhaitable de bénéficier d'une faible luminosité de l'ordre de 100 à 500 cd/m². En extérieur, il est souhaitable d'avoir une plus forte luminosité, de l'ordre de 2 000 à 5 000 cd/m².

Les deux sources vidéo les plus appropriées sont, du point de vue de la luminosité, la source LCoS et la source OLED. Notons cependant que la source OLED présente une luminosité limite pour des applications extérieures. Lorsque l'on regarde le paramètre de respect des couleurs, les sources OLED et CRT sont les plus adaptées. On peut donc conclure que la source OLED est à privilégier pour des applications en intérieur, pour son respect des couleurs et pour sa luminosité. En revanche en extérieur, il est préférable d'utiliser des sources LCoS.

Nous allons maintenant présenter un certain nombre de solutions optiques permettant d'afficher une image de la source vidéo devant l'utilisateur, via un dispositif transparent, ou

semi-transparent. Même s'il existe un certain nombre de solutions, utilisant différentes technologies optiques, les principes optiques mis en œuvre restent souvent les mêmes. Il est possible d'utiliser [158, 160]:

- Des lentilles et des lames semi-réfléchissantes
- Des lentilles et des prismes
- Des lentilles et dioptries asphériques
- Des lentilles et des éléments holographiques
- Des lentilles et des guides d'ondes
- Des lentilles et des guides d'ondes holographiques

Chacune, de ces techniques, présente des avantages et des inconvénients propres. Ces inconvénients peuvent être la taille du dispositif optique, la position par rapport à la tête, le poids du dispositif, la taille du champ visuel enrichi possible, la qualité des couleurs restituées et la présence de voile blanc ou coloré (voile provenant soit du rétroéclairage de la source vidéo, soit de couleurs résiduelles des traitements optiques utilisés). Cette liste n'est, bien sûr, pas exhaustive. Je vais présenter ici des solutions proches de celles proposées par la société Laster Technologies, car dans le cadre de mes travaux j'ai eu accès uniquement aux dispositifs de la société Laster Technologies [100, 99, 101, 102].

Pour caractériser un dispositif optique, il est important de s'intéresser à au moins 3 aspects :

- Le champ d'affichage : ce terme désigne la taille de l'image perçue par l'utilisateur. Cette valeur est exprimée en degrés. Il est soit caractérisé par la diagonale et le ratio entre la hauteur et la largeur, soit par les valeurs verticales et horizontales.
- La boîte de l'œil : ce terme désigne la taille de la zone dans laquelle l'utilisateur peut placer la pupille de son œil afin de voir l'image de la source vidéo. Cette valeur est exprimée en millimètre, et est définie en hauteur et largeur.
- La distance œil-élément : ce terme définit la distance entre l'œil et le premier élément du dispositif optique. Cette valeur est exprimée en millimètres. Cela permet de savoir s'il est possible d'utiliser une paire de lunettes de vue avec le système d'affichage.

Dans la famille des dispositifs d'affichage sur verres transparents, basés sur des éléments optiques réfractifs et semi-réfléchissants, nous pouvons trouver une pléthore de solutions se plaçant soit sur le côté de la tête soit sur le dessus. L'écran et les optiques relais sont alors soit placés sur le côté, soit au-dessus, soit en dessous de l'élément de mixage (qui mélange

l'image virtuelle à la réalité). Dans chaque dispositif optique, cf. Figure 5.5, il y a la source vidéo, un ensemble de lentilles pour mettre en forme le faisceau lumineux et au moins un élément semi-réfléchissant pour renvoyer l'image vers l'œil tout en permettant la vision de l'environnement. À titre illustratif, sur la Figure 5.5, nous avons :

- Un premier dispositif (à gauche): la source vidéo (42), complètement en haut sur la figure, puis un ensemble de 4 lentilles (43). L'élément semi-réfléchissant est une visière sphérique (44).
- Un second dispositif (à droite): la source vidéo (5) est en bas à gauche. Le faisceau lumineux est collimaté par un ensemble de 4 lentilles (4). Il est ensuite guidé par un prisme (10) et un guide d'onde (3), avant d'être envoyé vers l'œil.

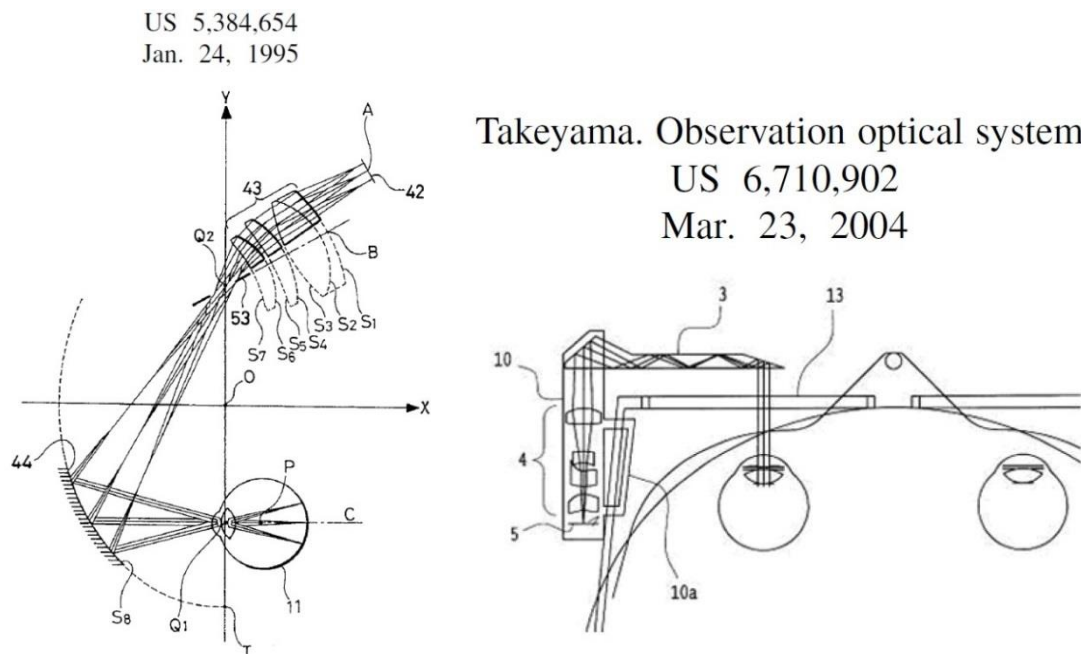


Figure 5.5 : Quelques exemples de dispositifs optiques de réalité augmentée [35].

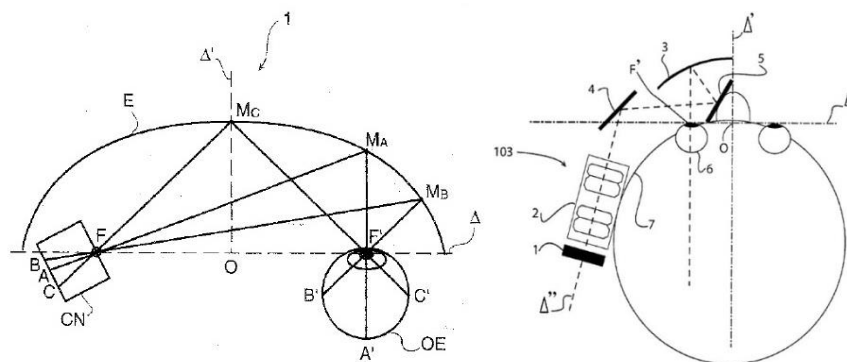


Figure 5.6 : Exemples de dispositifs de réalité augmentée de la société Laster Technologies [100, 99, 101, 102]

La Figure 5.6 représente quant à elle le principe utilisée par la technologie optique de l'entreprise Laster Technologies. La figure de gauche représente le principe général du dioptre asphérique : les points A', B' et C' de l'image rétinienne OE, sont les conjugués optiques des éléments A, B et C de la source vidéo CN, par le dioptre E (élément optique semi-réfléchissant). La figure de droite représente le dispositif réel : la source vidéo (1), les lentilles de focalisation (2), les miroirs de renvoi (4 et 5) et le dioptre asphérique semi-réfléchissant (3), permettant le renvoi du flux lumineux vers l'œil (6).

Ces dispositifs étant presque tous décentrés par rapport à l'axe de visée, ils présentent généralement des distorsions optiques (distorsion géométrique de l'image). Ces distorsions entraînent une variation de la densité de pixels, le nombre de pixels pour chaque zone de l'image varie en fonction de la distorsion (comme sur la Figure 5.7), pour chaque zone du champ visuel, et de la luminosité. En effet, tandis que certaines zones de l'image sont comprimées par le système optique, d'autres zones sont soit non modifiées soit étirées. Ces variations entraînent une variation du nombre de pixels par unité de surface du champ visuel.

Lorsque l'on compense une distorsion optique par traitement d'image, l'image est déformée afin de compenser cette distorsion, cela impacte la densité de pixels, cela a par conséquent un impact sur la qualité de l'image finale. Comme le montre la Figure 5.7, pour obtenir une image non distordue à travers un système optique hors d'axe, il est souvent nécessaire de pré déformer l'image. Lorsque l'on affiche une image pré déformée sur la source vidéo, cf. image du milieu, il est possible d'obtenir l'image souhaitée sans distorsion à travers le dispositif d'affichage, cf. image de droite. Cependant, comme l'image à afficher sur la source vidéo est déformée, celle-ci n'a pas la même densité de pixels pour toutes les zones de l'image. Il est souhaité une image ayant des barres verticales de même taille et donc si possible ayant le même nombre de pixels. Or la distorsion géométrique due à l'optique implique une variation du nombre de pixels de la source vidéo par unité angulaire du champ visuel.

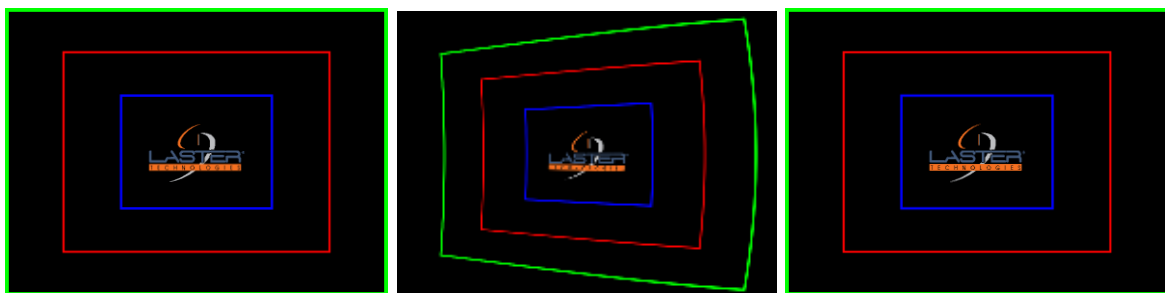


Figure 5.7 : De gauche à droite : l'image originale, l'image incluant une contre distorsion à afficher sur la source vidéo et l'image observée par l'utilisateur à travers le système.

Il est important de pouvoir corriger à la fois les distorsions géométriques [159] de l'image et les variations de colorimétrie. Pour corriger les variations colorimétriques, il est possible de se baser sur les méthodes employées pour la projection cinématographique [64]. Il s'agit généralement d'aligner le gamut couleur [25, 117, 33] disponible à travers le système à celui souhaité.

5.1.2.2. Monoculaire vs binoculaire

En vision monoculaire, un seul œil a une vision enrichie (vision de la réalité et de l'image fournie par le dispositif d'affichage) ce qui peut générer des disparités entre les deux images rétiniennes, cf. Figure 5.8. Dans ce cas, afin de voir correctement l'image affichée sur l'écran en monoculaire, il faut afficher l'image dans le même plan que le plan de focalisation de l'œil. Autrement dit, il faut que l'observateur puisse voir le monde réel net et l'image affichée avec la même accommodation.

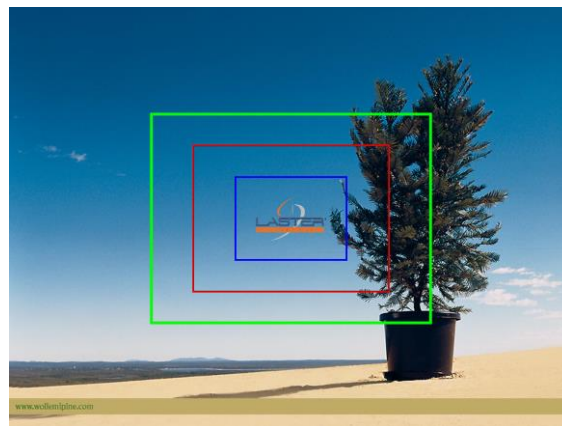


Figure 5.8 : Image monoculaire pour un champ de vision augmenté de 40 X 30 degrés et un champ de vision total de 80 X 60 degrés.

En vision binoculaire, il faut que les images affichées correspondent afin de ne pas perturber la vision binoculaire. En vision monoculaire, il y a une rivalité entre les deux images rétiniennes (une seule est enrichie), par contre il n'y a pas de problème lié à la relation convergence-accommodation [50]. En vision binoculaire, l'observateur doit maintenir la relation accommodation-convergence, il faut donc que les images affichées soient calculées convenablement pour ne pas perturber la fusion des images rétiniennes [137].

Les dispositifs monoculaires présentent l'avantage de ne pas nécessiter la génération d'images binoculaires. Ils peuvent être mis en œuvre facilement sur l'œil dominant de la personne, afin de minimiser la rivalité entre l'œil équipé et l'œil non équipé [50].

Les dispositifs binoculaires présentent comme intérêt l'absence de rivalité entre les deux yeux. Cependant, cela impose la génération d'images binoculaires confortables. Dans le cas d'assistance aux personnes malvoyantes, les aspects de vision simultanée et de fusion d'images sont importants. Il est donc important de prendre en compte d'une part la correction optique de la personne afin de lui permettre de bénéficier de son accommodation et de sa convergence de manière convenable et sans contrainte [191] et d'autre part de prendre en compte la convergence, l'accommodation et les écarts pupillaires [184] de la personne afin de conserver la relation accommodation-convergence.

Un dispositif de réalité augmentée binoculaire à verres transparents présente l'avantage de pouvoir utiliser soit l'affichage sur l'œil gauche, soit sur l'œil droit, soit sur les deux yeux, en fonction de la situation, et de l'état de fatigue du malvoyant.

Les images ci-dessous, cf. Figure 5.9 et Figure 5.10, présentent deux cas de vision binoculaire : le cas confortable et le cas inconfortable. La Figure 5.9 représente une même scène, 80 X 60°, et la perception de l'image binoculaire, de 60 X 30°, pour un observateur, la partie de gauche représente le cas confortable et la partie de droite le cas inconfortable. La Figure 5.10 représente les images gauches et droites, à afficher sur les modules optiques, permettant d'obtenir les situations de la Figure 5.9, à gauche le cas confortable et à droite le cas inconfortable. La configuration illustrée correspond à une rotation, vers l'extérieur de chaque module optique, d'environ 8°, offrant un champ de vision augmenté de 60 X 30 degrés. Le cas de droite présente une situation où les images sont mal générées pour le dispositif binoculaire, le dédoublement des barres verticales bleu, ne devrait pas exister.

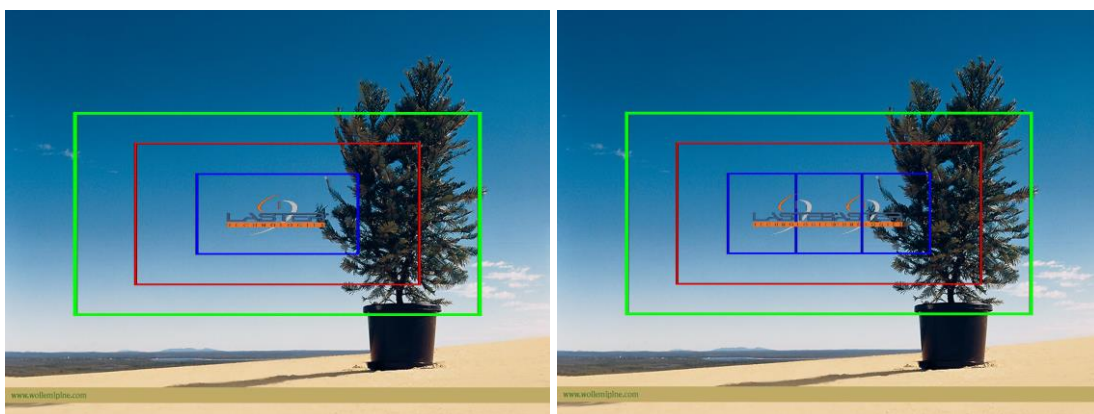


Figure 5.9 : Image finale binoculaire superposées à une scène : à gauche la situation confortable et à droite une situation inconfortable

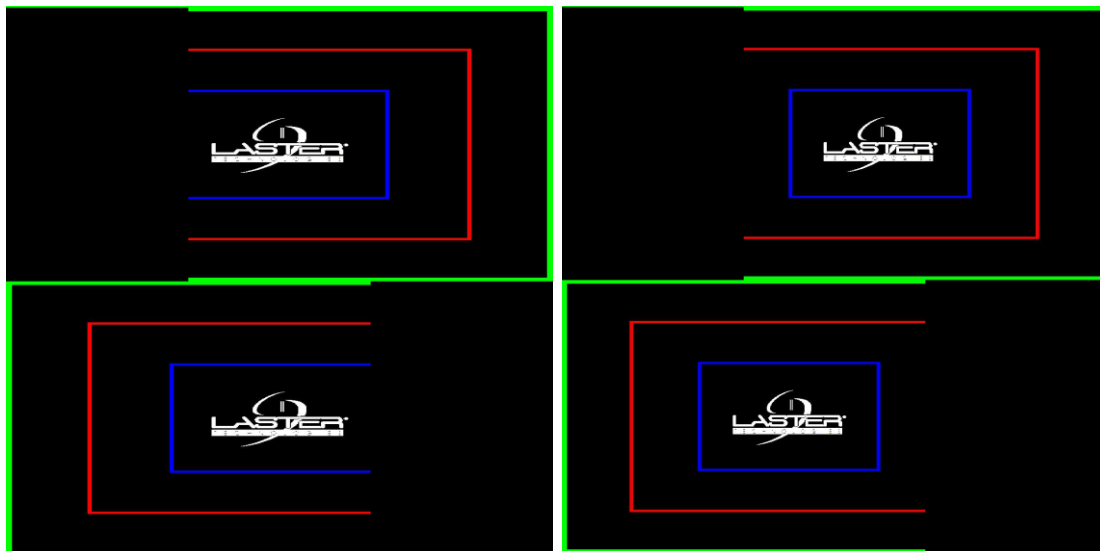


Figure 5.10 : Images gauche et droite pour une situation confortable (colonne de gauche) et inconfortable (colonne de droite).

5.1.2.3. Modèle de génération d'image binoculaire

Pour pouvoir afficher une image binoculaire confortable, j'ai développé un modèle de génération d'image adaptée qui prend en compte les spécificités des dispositifs optique d'affichage de réalité augmentée.

Je vais maintenant présenter le modèle, le processus, que j'ai développé, pour générer des images binoculaires confortables en fonction du dispositif optique d'affichage à verres transparents utilisé. Ce processus utilise la modélisation du champ optique, de la position de l'écran et de l'image virtuelle du dispositif dans l'espace [176]. Ce processus prend en compte la position relative d'un module par rapport à l'autre (module optique gauche et module optique droit), en prenant en référence la position de l'œil cyclope, œil virtuel placé au milieu sur la ligne reliant les deux yeux de la personne, l'objectif étant de calculer les points de fusion entre les images gauches et droites. Une fois la position de chaque module connu, il est possible de déterminer la taille et la forme de l'image binoculaire. Une fois cette image définie, il est possible de calculer, par rapport à la position de chaque module optique, l'image déformée à afficher pour le module gauche et pour le module droit.

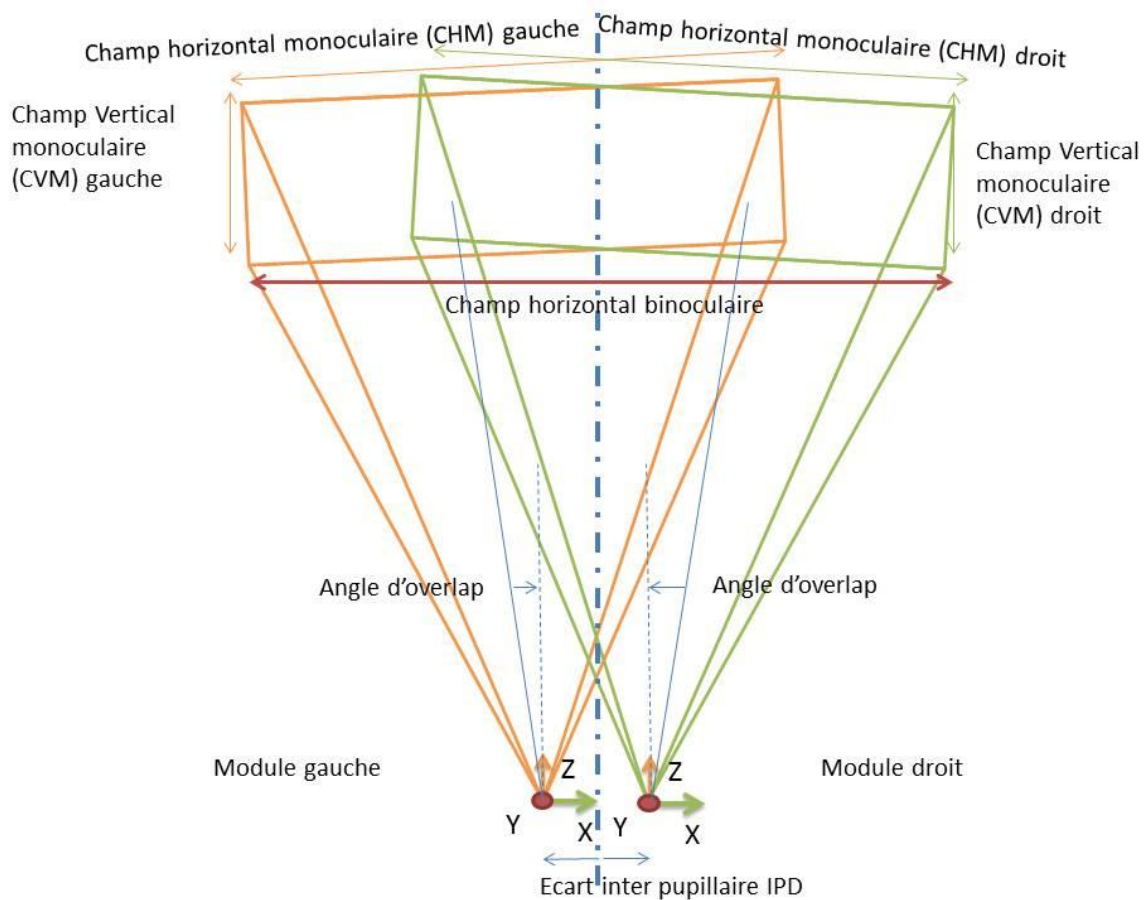


Figure 5.11 : Schéma du modèle binoculaire

Comme le montre la Figure 5.11, il est possible de modéliser, représenter de manière schématique, un dispositif binoculaire en fonction de la position des deux modules optiques monoculaires (gauche et droit). Le module de l'œil gauche est ici représenté en orange, celui de l'œil droit en vert.

Il est possible en définissant la distance de focalisation des deux modules optiques, la distance inter-pupillaire (IPD) théorique du système (assimilée à la distance mécanique entre le centre des deux boîtes de l'œil), les champs de vision monoculaire verticaux et horizontaux de chaque module, de générer les images gauches et droites correspondant à une seule image binoculaire confortable. Afin d'élargir le champ de vision binoculaire, il est possible de tourner les deux modules vers les tempes (vers l'extérieur du dispositif), l'angle ainsi appliqué aux modules optiques est appelé par la suite « angle d'overlap ».

Il est alors possible de définir la position de l'image binoculaire en posant :

- $X=0$,

- $Y=0$,
- Z = distance de focalisation DF ,
- sa largeur L et sa hauteur H du champ d'affichage [50, 176].

Alors la largeur de l'image binoculaire vaut:

$$L = 2 * DF * \tan\left(\frac{CHM}{2} + \text{Overlap}\right) + IPD$$

Avec CHM = champ horizontal du module monoculaire

Le champ vertical de l'image binoculaire vaut la valeur du champ vertical monoculaire du module. La hauteur de l'image binoculaire vaut :

$$H = 2 * DF * \tan\left(\frac{CVM}{2}\right)$$

Avec CVM = champ vertical du module monoculaire

On peut alors exprimer la position du module optique gauche par :

- $X = \frac{-IPD}{2}$,
- $Y = 0$,
- $Z = 0$

Et les rotations suivant les axes :

- $X = 0$,
- Y = -l'angle d'overlap,
- $Z = 0$.

De même, la position du module droit est définie par :

- $X = \frac{IPD}{2}$,
- $Y = 0$,
- $Z = 0$

Et les rotations suivant les axes :

- $X = 0$,
- Y = l'angle d'overlap,

- $Z = 0$.

Lorsque les paramètres sont corrects (par rapport aux modules optiques gauches et droits et correspondent aux paramètres mécaniques), il est alors possible d'obtenir des images binoculaires confortables, comme le montre les Figures 5.12 et 5.13.

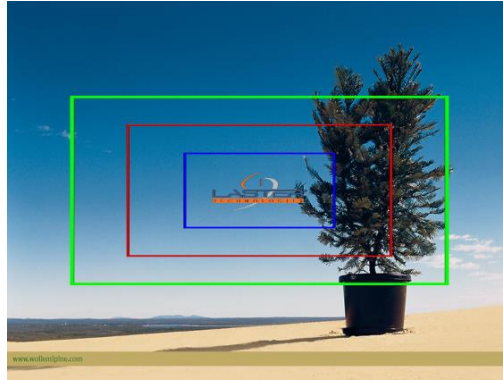


Figure 5.12 : Image finale binoculaire superposée à une scène observée

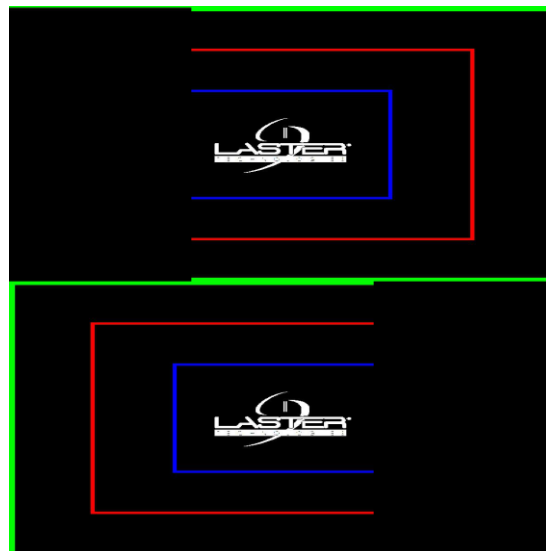


Figure 5.13 : Images gauche et droite correspondant à une situation confortable du type Figure 5.12

5.1.3. Recalage des informations

Dans cette partie, je vais présenter la problématique du recalage des informations en réalité augmentée du point de vue technique. Pour cela, je vais présenter dans un premier temps la problématique du recalage et recenser des solutions possibles.

Les données à afficher sont référencées par rapport au repère du capteur ayant permis leur acquisition et leur localisation dans l'espace. Le dispositif d'afficheur possède lui aussi son

propre repère. La différence entre ces repères génère, si elle n'est pas compensée, des décalages entre la position réelle et la position souhaitée pour les informations à afficher [108]. Il est donc nécessaire de connaître tous les paramètres décrits précédemment pour chaque élément du dispositif de réalité augmentée. Pour un dispositif d'affichage, cela implique notamment de connaître la distorsion optique du dispositif. Il est également nécessaire de connaître la position de chaque capteur par rapport au dispositif d'affichage [128] afin de compenser ces décalages [157], ainsi placer au bon endroit les informations à afficher, cf. Figure 5.14, afin de réduire à son minimum les erreurs de positionnement des informations [82].

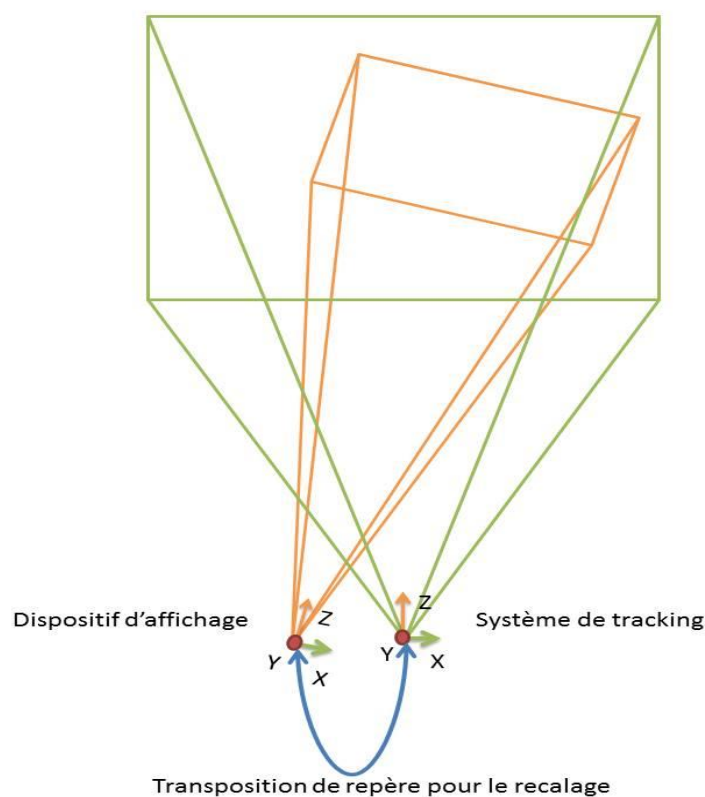


Figure 5.14 : Chaîne de calibration de réalité augmentée

Il est indispensable de connaître les distorsions réelles du dispositif d'affichage, des capteurs et leurs champs respectifs réels afin de pouvoir utiliser ces informations au lieu des données théoriques, cela permet d'obtenir de meilleurs résultats pour l'affichage des informations enrichies par rapport à la réalité. Il est alors possible, comme indiqué précédemment, de transposer les positions des informations à afficher du référentiel du capteur à celui du dispositif d'affichage.

Une fois ce calcul réalisé, il est possible de calculer l'échelle et la position exacte des informations à afficher. La Figure 5.15 montre deux exemples, de scène de réalité augmentée, réalisés avec un dispositif de la société Laster technologies, pour lesquels on affiche un parallélépipède placé sur un damier noir et blanc.

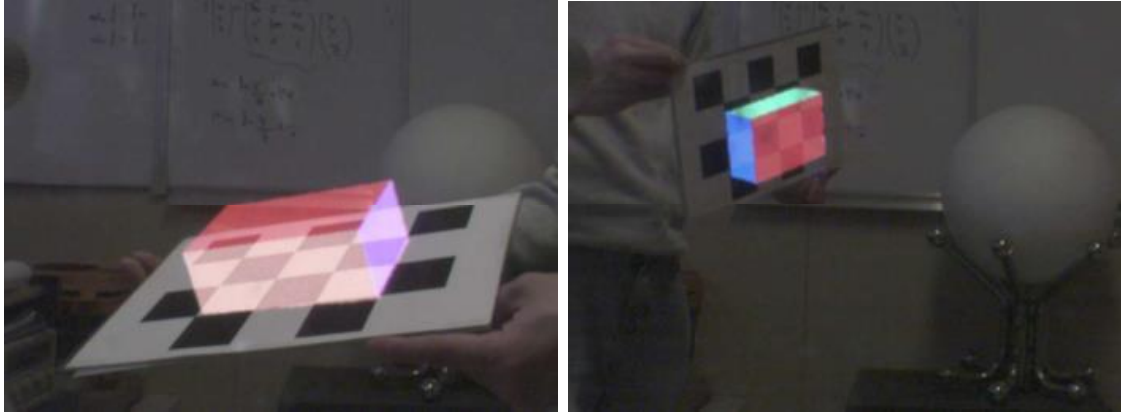


Figure 5.15 Illustrations du recalage d'information en réalité augmentée

Sur l'image de gauche de la Figure 5.15, le parallélépipède est correctement placé sur le damier noir et blanc (il est correctement centré dessus). En revanche, sur l'image de droite, ce parallélépipède est légèrement trop bas (il n'est pas tout à fait centré).

Afin de développer une solution algorithmique qui prend en compte ces erreurs et les compense en temps réel, ces erreurs ont été analysées et mesurées afin de calibrer le système [128, 76].

Actuellement, nous disposons uniquement de l'information de contours 2D, pour le détecteur de contours présenté au chapitre 3 ; afin d'effectuer le recalage de ces informations j'utilise des offsets et des mises à l'échelle de notre carte de contours. Cette technique sera présentée dans la 3^{ème} partie de ce chapitre, lors de la présentation du premier démonstrateur réalisé.

5.1.4. Choix du dispositif d'affichage

Les dispositifs de réalité virtuelle et de réalité augmentée présentent chacun des avantages et inconvénients [103]. Alors que, l'aspect immersif est préférable pour certaines situations, pour d'autres situations telles que l'appréciation des distances [90] et l'assistance aux personnes malvoyantes en mobilité, les systèmes à verres transparents [81] sont préférables car le malvoyant conserve la perception de la réalité dans laquelle il évolue.

Pour l'usage d'un tel dispositif dans des conditions optimales, il est très important que le design de celui-ci présente certaines caractéristiques, à savoir une image virtuelle confortable à observer pour un usage prolongé, un système ouvert sur la réalité, mécanique présentant peu d'éléments obstruant le champ de vision [54, 195].

Comme nous l'avons montré dans ce début du 5^{ème} chapitre, un système binoculaire permet de fournir un grand champ de vision [103] sans rivalité binoculaire (opposition entre une seule image virtuelle et la vision binoculaire de la réalité), c'est pour ces raisons que j'ai retenu ce type de dispositif pour mes développements.

5.2. Système d'acquisition : caméra

Pour réaliser une application de réalité augmentée sur la base d'un dispositif d'affichage à verre transparent, il est nécessaire de disposer d'informations extraites d'un ou plusieurs capteurs tels que les centrales inertielle ou les caméras. Il faut alors pouvoir calculer, analyser, ces images dans un temps suffisamment court et afin d'afficher correctement ces informations. Nous allons maintenant analyser les paramètres déterminants dans le choix d'une caméra pour un dispositif de réalité augmentée.

5.2.1. Introduction

Les capteurs jouent un rôle important dans un dispositif de réalité augmentée. Pour passer de la vision augmentée à la réalité augmentée, il est nécessaire de pouvoir analyser l'environnement en temps réel. Il est donc nécessaire de disposer de capteurs tels que les caméras, les centrales inertielle ou les capteurs GPS.

Dans la tâche qui nous concerne, à savoir l'assistance visuelle aux personnes malvoyantes, il est intéressant de disposer d'une image de la réalité en temps réel. Pour cela, je me suis intéressé aux différentes technologies de capteurs imageurs. Je me suis uniquement focalisé sur les caméras matricielles, compte tenu de l'application visée.

Nous allons commencer par présenter le fonctionnement, les avantages et inconvénients des capteurs CCD, puis nous analyserons les capteurs CMOS. Nous n'allons pas présenter ici tous les paramètres liés aux caméras, mais uniquement les paramètres les plus utiles pour une application de réalité augmentée.

5.2.2. Caméra CCD

Une caméra est composée d'éléments photosensibles. Ils sont généralement disposés en matrice comprenant un certain nombre de lignes et de colonnes. Le nombre de lignes et de colonnes définit, la résolution de la caméra. Pour chaque case, ou élément de la matrice, il y a un élément photosensible communément appelé « pixel » [181].

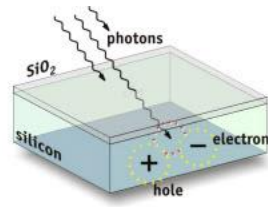


Figure 5.16 : Schéma de principe d'un élément photosensible⁶¹

Comme le montre la Figure 5.16, les éléments photosensibles vont être sensibles à la lumière, aux photons. Le pixel va convertir ces photons en électrons, on parle alors de charge. Le terme CCD vient de cette collecte et conversion des photons en potentiel électrique, ou charge. Lorsque l'on réalise une image à partir de ce type de capteurs, tous les éléments photosensibles se remplissent en même temps en électrons. Une fois la fin du temps d'exposition du capteur (temps pendant lequel les éléments photosensibles se chargent) la charge de chacun des pixels est transférée, cf. Figure 5.17, jusqu'à la sortie du capteur. Généralement chaque colonne de pixels est transférée à une colonne de transfert, puis chaque colonne est transférée en mémoire [181].

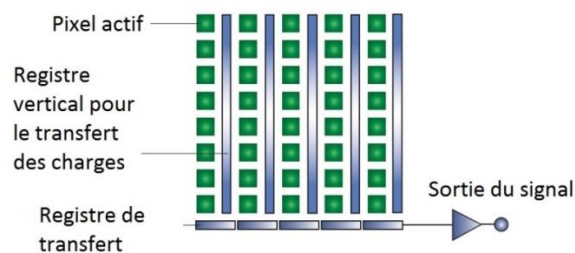


Figure 5.17 : Transfert des charges dans une matrice CCD

Ce principe de fonctionnement présente l'avantage que tous les pixels ont acquis les charges au même moment, il n'y a donc pas d'effet lié au décalage temporel des informations acquises par les pixels. Les cellules photosensibles utilisées dans les matrices CCD peuvent dans certains cas saturer. Dans ce cas, les charges acquises par les pixels concernés peuvent être

⁶¹ http://www.teledynedalsa.com/img/imaging/photoelectric_210w.jpg

transmises par erreur (comme par débordement) sur les pixels voisins. Dans ce cas, il apparaît des « trainées » appelées « Blooming », en premier lieu suivant une des directions de parcours de la matrice, généralement celle du registre vertical. Dans certains cas, il peut y avoir un débordement suivant les deux directions de la matrice, cf. Figure 5.18.

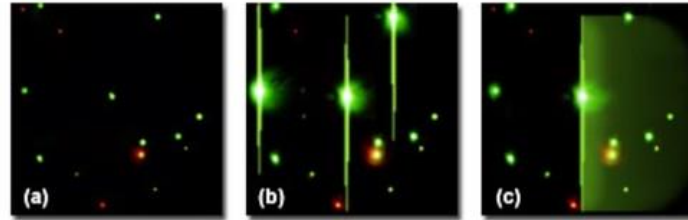


Figure 5.18 : (à gauche) Image sans blooming ; (au centre) avec un blooming vertical ; (à droite) avec un blooming vertical et horizontal.⁶²

Ce phénomène de blooming rend difficilement exploitable les pixels saturés et les pixels voisins qui de fait se trouvent affectés ce qui impacte les algorithmes de traitement d'image. Ces artefacts créés par le capteur peuvent rendre inutilisable tout ou partie de l'image pour aider les personnes malvoyantes.

5.2.3. Caméra CMOS

Il existe également des caméras à base de matrice CMOS⁶³ (généralement moins chères, que les matrices CCD). Ces matrices sont constituées d'éléments mesurant l'énergie lumineuse arrivant dessus [181] et non au travers de puits de lumière comme les CCD.

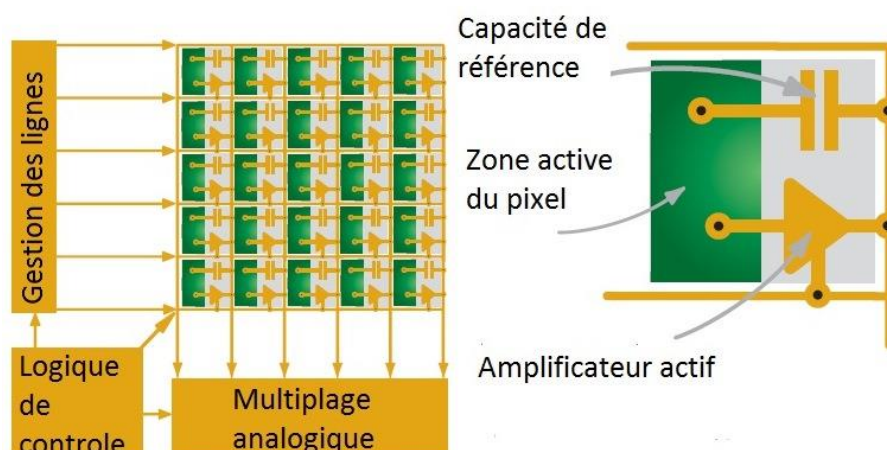


Figure 5.19 : Schéma de principe d'une matrice CMOS

⁶² http://www.optique-ingenieur.org/fr/cours/OPI_fr_M05_C06/co/Contenu_12.html

⁶³ http://www.optique-ingenieur.org/fr/cours/OPI_fr_M05_C06/co/Contenu_18.html

Ces matrices présentent certains avantages tels que l'adressage séparé de chaque pixel. De plus, en cas de saturation, les pixels de type CMOS ne débordent pas sur leurs pixels voisins, seul le signal du pixel concerné est saturé. Ces différences sont dues à une gestion différente des pixels au sein de la matrice, cf. Figure 5.19.

Les matrices CMOS présentent des avantages au niveau de la rapidité de lecture des images et de l'absence de blooming. Cependant, il existe aussi des inconvénients : un bruit plus élevé que pour les matrices CCD et une variation de la sensibilité entre les pixels plus élevée également⁶⁴.

Les matrices CMOS nécessitent la présence d'un « shutter » ou obturateur électronique, il s'agit d'un mécanisme qui gère le temps d'exposition de chaque pixel. Il existe deux types d'obturateurs électroniques : « Rolling shutter » et « global shutter ». Dans un premier temps, je vais présenter ces deux types de fonctionnement, leurs avantages, et inconvénients. Il existe dans certains cas quelques pistes pour compenser ces inconvénients et ainsi améliorer la qualité de l'image. Dans un deuxième temps, je vais expliquer mon choix pour notre application de réalité augmentée.

Il existe plusieurs types de rolling shutter. Nous n'entrerons pas dans les détails de chacun d'entre eux, nous allons seulement présenter le principe général. Les matrices CMOS à rolling shutter constituent le type de capteur d'images le plus répandu dans les appareils grand public tels que les consoles de jeux, les téléphones portables, les smartphones, les ordinateurs portables, les webcams et les interphones vidéo. De manière générale, chaque pixel image est parcouru de manière séquentielle, un pixel après l'autre. L'acquisition de chaque pixel se termine juste avant la lecture de celui-ci. L'instant capturé par chaque pixel est donc légèrement différent pour chacun des pixels de l'image ainsi obtenue. Ce phénomène entraîne un décalage temporel en chaque point de l'image, qui génère un effet de mouvement dans l'image alors qu'aucun mouvement de ce type n'est présent dans la scène [60, 61]. Cela peut déformer les objets, cf. Figure 5.20, et diminuer la performance de certains algorithmes de traitement d'images, par exemple la détection de lignes ou de formes géométriques. Ces variations temporelles peuvent également se conjuguer à des variations photométriques, issues d'une variation de l'illumination de la scène. Sur la Figure 5.20, l'image acquise par le capteur « rolling shutter » (à gauche), chaque pixel étant acquis à un instant légèrement différent, la rotation des pales du ventilateur génère une image peu ressemblante à l'objet réel.

⁶⁴ http://www.optique-ingenieur.org/fr/cours/OPI_fr_M05_C06/co/Contenu_20.html

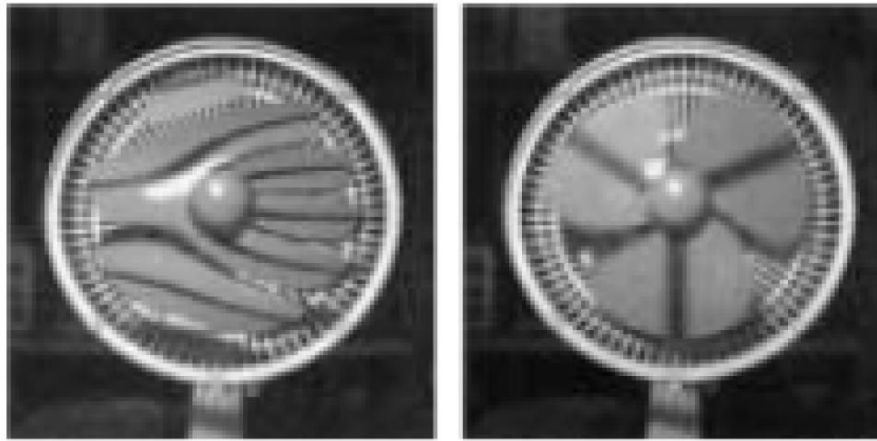


Figure 5.20 : (à gauche) Image acquise par une caméra CMOS « rolling shutter » ; (à droite) une caméra « global shutter »

Il est également possible de réaliser des capteurs CMOS « global shutter ». Dans ce cas, le « rolling shutter » du capteur est remplacé par un shutter dit global. Pour cela, il est généralement ajouté au capteur une mémoire qui stocke l'image, puis la transfère pendant l'acquisition de l'image suivante. Cela permet de bénéficier des avantages de la technologie CMOS (avec un prix proche des capteurs CMOS rolling shutter) sans l'inconvénient du « rolling shutter ». Cela nécessite de synchroniser électroniquement tous les pixels lors de la phase d'acquisition de l'image, puis lors de leur transfert dans la mémoire. Il est alors possible de transférer l'image en une seule passe sans l'inconvénient de l'adressage de chaque pixel les uns après les autres. Tous les pixels sont alors sensibles au même instant pour réaliser en même temps la capture de la scène, ce qui permet d'obtenir une image sans les artefacts liés au « rolling shutter ».

5.2.4. Critères de sélection

Comme nous venons de le voir, il existe plusieurs types de capteurs pour acquérir des images. Les capteurs CCD ne sont pas très adaptés pour l'acquisition d'images dans un environnement extérieur non contrôlé. Avec ce type de capteur, il est impossible de pouvoir prédire et éviter les saturations et sur-illuminations dans la scène et par conséquent d'éviter le phénomène de « blooming ». Il est préférable dans ce contexte de mobilité d'utiliser un capteur de type CMOS de type « global shutter ». Rappelons que pour notre application de réalité augmentée, nous devons d'une part disposer d'images non saturées, et d'autre part disposer d'images fidèles à la scène, c'est-à-dire non impactées par des artefacts et des

mouvements induits par l'obturateur du capteur. Dans le cadre de notre application, nous utiliserons donc un capteur de type CMOS « global shutter ».

Le type de capteur et la technologie de l'obturateur sont très importants pour notre application, cela étant, on doit également prendre en compte un certain nombre d'autres paramètres dont :

- la résolution : il s'agit du nombre de pixels disponible sur le capteur. Cette valeur est indiquée en nombre de pixels en hauteur et en largeur sur la matrice. Plus un capteur aura de pixels, plus il sera possible de discerner des détails fins dans les images, cependant cela entraîne un nombre beaucoup plus grand de données à transmettre et à traiter. Au niveau des CCD et CMOS destinées aux applications professionnelles, il existe aussi bien des caméras de faibles résolutions (inférieur à 320 X 240) que de grandes résolutions telles que 2560 X 2048. Pour les caméras miniatures destinées aux produits grand public tels que les smartphones, il existe des caméras disposant également de 18 et 42 millions de pixels.
- le temps d'exposition : il s'agit du temps pendant lequel, les pixels de la matrice sont exposés à la lumière provenant de la scène à acquérir. Plus ce temps est long, plus l'image sera lumineuse. Plus ce temps est court plus l'image sera sombre. En cas de scène mobile, il est primordial d'utiliser un temps d'exposition court afin de ne pas créer d'artefacts dus aux déplacements des objets pendant l'acquisition de l'image.
- la cadence : il s'agit du nombre d'images que la caméra peut réaliser par seconde. Cette valeur est exprimée en nombre de trames par seconde « FPS » (frame per second). Plus la taille de l'image est grande moins la cadence peut être élevée. Pour des applications de réalité augmentée, il est souvent nécessaire de disposer d'au moins 15 voire 30 FPS. Si la cadence est trop faible, lorsque l'image est analysée les informations extraites sont affichées trop tard, en décalé par rapport à la perception du monde réel.
- la sensibilité : il s'agit de la plus petite illumination à partir de laquelle la caméra est capable de fournir une image exploitable. L'image obtenue contient alors un signal qui n'est pas uniquement du au bruit (variation de la réponse de chaque pixel, les uns par rapport aux autres pour une même illumination).
- la dynamique : cette valeur désigne la capacité de la caméra à restituer des informations très sombres ou très lumineuses de manière correcte. Cela représente la

plage du signal en sortie de la caméra qui est réellement utilisable pour analyser la scène. Cela prend en compte le bruit du capteur.

- la taille et le poids : ces paramètres sont importants pour l'intégration de la caméra dans un dispositif d'affichage. Beaucoup de caméras professionnelles ont un volume supérieur à un cube de 30 mm de côté et un poids supérieur à 10 g. Les caméras grand public, elles ont quant à elle des tailles de 6 à 10 mm de côté et entre 2 et 6 mm d'épaisseur pour un poids de quelques grammes.

Au vu de ces différents paramètres, j'ai réalisé un grand nombre de tests de caméras afin de trouver un capteur répondant aux critères de mon application à savoir : CMOS, global shutter, résolution moyenne VGA (640 X 480 pixels en couleurs), 30 FPS, usage en intérieur et extérieur, poids inférieur à 10g et pouvant être intégré dans une paire de lunettes. J'ai notamment choisi et testé le capteur Aptina MT9V024 qui est un capteur CMOS, global shutter fournissant 752 X 480 pixels en couleurs, 60 FPS, sensible à partir de 1 Lux et disposant d'un mode HDR (high dynamic range), ce mode permet d'obtenir des images en intérieur et extérieur sans avoir à modifier les réglages de la caméra. Cette caméra peut être modifiée en une carte sur-mesure avec un port USB2.0 ou USB3.0 et pour des dimensions compatibles avec l'intégration dans une paire de lunettes de réalité augmentée, soit environ 50 X 15 mm.

Toute caméra nécessite un objectif pour focaliser la scène à acquérir sur le capteur. Il existe une pléthore de modèles sur le marché. Le choix d'un objectif doit permettre une restitution fidèle des couleurs, si possible peu de distorsion optique et être compact pour ne pas impacter le design de la paire de lunettes.

5.3. Unité de calcul

Dans cette dernière partie, je vais présenter les implémentations logicielles que j'ai réalisées basées sur ce chapitre et les chapitres précédents. Je vais également présenter le premier démonstrateur que j'ai développé. Celui-ci intègre à la fois une caméra analogique « global shutter » miniature, un dispositif d'affichage monoculaire et un logiciel de traitement d'images et d'affichage.

Pour des raisons techniques, j'ai réalisé ce premier démonstrateur sur la base d'une paire de lunettes monoculaire. La partie binoculaire sera bientôt intégrée lors d'une évolution du logiciel et du matériel. Plus précisément, j'ai utilisé une paire de lunettes monoculaire

produite par Laster Technologies, à savoir le produit EyePhone Professionnal® (EPV Pro). Ce dispositif présente l'intérêt de ne pas nécessiter de contre distorsion, c'est-à-dire qu'il n'est pas nécessaire de déformer les images afin d'obtenir des images correctes et non distordues à travers le dispositif optique. Enfin, ce dispositif ne nécessite pas de réglage d'écart pupillaire.

5.3.1. Implémentation logicielle

Dans un premier temps, je vais présenter l'architecture générale du logiciel mis en œuvre, puis présenter les performances obtenues.

Afin de prendre en compte le champ utile de la caméra (afin d'appliquer les traitements d'images uniquement sur la zone intéressante de l'image à traiter), j'ai défini cette zone par les coordonnées de son coin supérieur gauche et en fonction de sa hauteur et sa largeur, cf. Figure 5.21. Ces informations m'ont permis de conserver uniquement la zone utile de l'image. Cette zone mesure 220 pixels par 240 avec la caméra disponible pour mes tests.

The image shows a software interface with two main sections for adjusting regions of interest. At the top, there are tabs for 'Users', 'Magnifier settings', and 'Inputs / outputs'. The first section, 'Input area from image', contains four input fields: 'X Value' (0), 'Y Value' (0), 'Height Value' (480), and 'Width Value' (640). Each field has 'minus' and 'more' buttons for adjustment. The second section, 'EyePhone Video Pro useful area', also contains four input fields: 'X Value' (0), 'Y Value' (0), 'Height Value' (600), and 'Width Value' (800), each with 'minus' and 'more' buttons.

Figure 5.21 : Interface pour l'ajustement des zones d'intérêts caméra et écran

De même, pour afficher les informations uniquement dans la zone utile l'écran, j'ai procédé de même pour définir la zone utile pour l'image à afficher, cf. Figure 5.21. Cette zone mesure environ 520 X 520 pixels. L'image envoyée à l'écran est une image de fond noir de résolution 800 X 600 pixels. Seule la zone utile de l'image à afficher contient des informations couleurs. Afin d'afficher l'image de la caméra, dans de bonnes proportions, j'ai ajusté la taille de l'image acquise à la taille de l'image à afficher (dans mon cas 520 X 520 pixels). Ensuite j'ai inséré dans l'image noire (image écran de 800 X 600 pixels), la zone de 520 X 520 pixels, correspondant l'image à restituer.

N'ayant pas accès à la localisation dans l'espace de chaque contour, mais uniquement d'une information bidimensionnelle, il n'est pas possible de réaliser un recalage précis tel que présenté à la section 5.1.3. J'ai contourné ce problème en réalisant un recalage en 2 dimensions des contours pour une distance donnée. Dans ce cas, il faut redimensionner l'image à afficher afin de lui donner la bonne échelle puis de la positionner au bon endroit en appliquant des offsets verticaux et horizontaux. Ce recalage est valable pour une distance donnée. Ce principe est utilisable à plusieurs distances à condition de pouvoir disposer d'un recalage précis pour chacune des distances et un moyen de sélectionner celle-ci. Dans ce cas, il peut même être possible de proposer des réglages différents en fonction de la distance. En effet, il peut être utilisé des réglages (taille de rehaussement, couleurs) différents en fonction que l'on soit en vision de près ou de loin.

La Figure 5.22 représente le principe de fonctionnement de l'application.

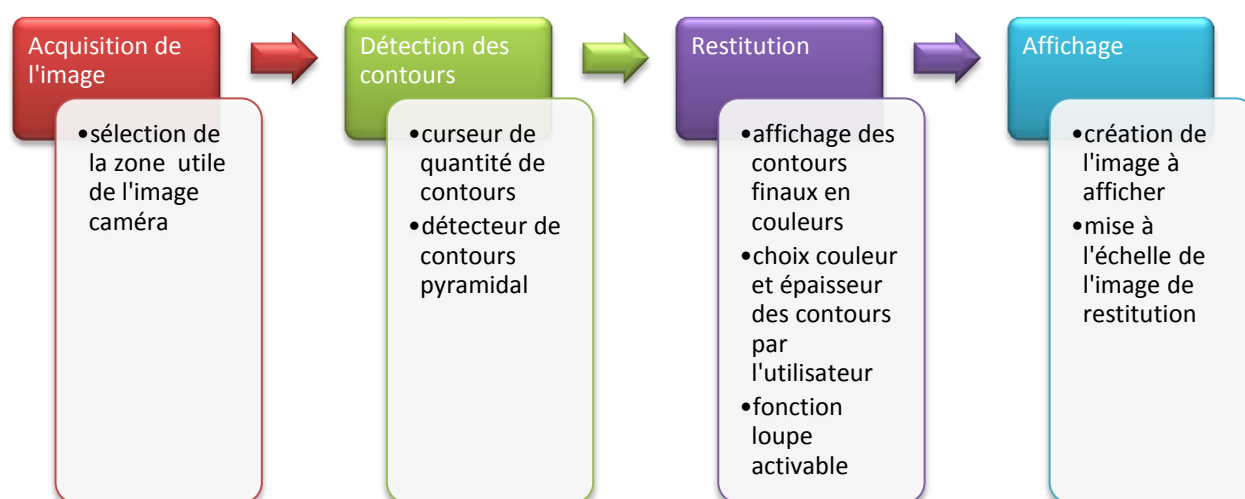


Figure 5.22 : Schéma général de l'application

Après avoir défini la partie acquisition et affichage, je vais maintenant présenter la partie extraction des informations. Comme vu au chapitre 2, il est nécessaire que le malvoyant interagisse avec le dispositif. Pour cela, j'ai développé au chapitre 3 une méthode d'extraction de contours pour laquelle il est possible d'indiquer la « quantité de contours souhaitée » (aussi appelé « gaussian filter size »), cf. Figure 5.23, et de choisir la couleur (colors restitution mode) et la taille (width of mark for restitution) des contours. Ces paramètres permettent à l'utilisateur de définir « la quantité de contours souhaitée » et de pouvoir adapter en temps réel la manière de restituer ces contours dans son champ visuel.

Users : Magnifier settings : Inputs / outputs :

User :

Gaussian filter size : 1

Width of mark for restitution : 0

Color restitution mode : Yellow

Select Magnifier mode : Edges

Select magnifier factor : 1

Select color for magnifier : White

Figure 5.23 : Paramètres utilisateurs

Comme vu au chapitre 2, les personnes malvoyantes ont parfois besoin, lors de leurs déplacements de disposer d'une loupe ou d'une fonction similaire. J'ai également évoqué qu'une loupe portée sur tête présente l'avantage de rester alignée avec l'axe du regard. En effet, l'alignement de l'œil avec ce type de dispositif est une contrainte importante pour son usage visé. Dans notre cas, j'ai donc intégré, pour simplifier cette tâche, une fonction loupe à notre dispositif. L'utilisateur n'a alors qu'à placer l'objet à agrandir au centre de son champ visuel, cf. Figure 5.25. Sur l'interface de la Figure 5.23 j'ai donc intégré 3 paramètres qui permettent d'ajuster le grossissement de la loupe (magnifier factor), la couleur de la loupe (color for magnifier) et le mode (magnifier mode). Ce dernier permet de choisir le mode d'affichage. Il est, en effet, possible d'afficher l'image d'origine uniquement, les contours uniquement, la loupe uniquement, ou la loupe sur une partie de l'écran avec ou sans les contours. Afin d'afficher la loupe, j'ai défini la zone du champ d'affichage sur laquelle elle doit être, cf. Figure 5.24.

Magnifier position :

X Value: minus 0 more Y Value: minus 0 more

Height Value: minus 100 more Width Value: minus 100 more

Figure 5.24 : Réglage de la position relative de la loupe dans la zone d'affichage

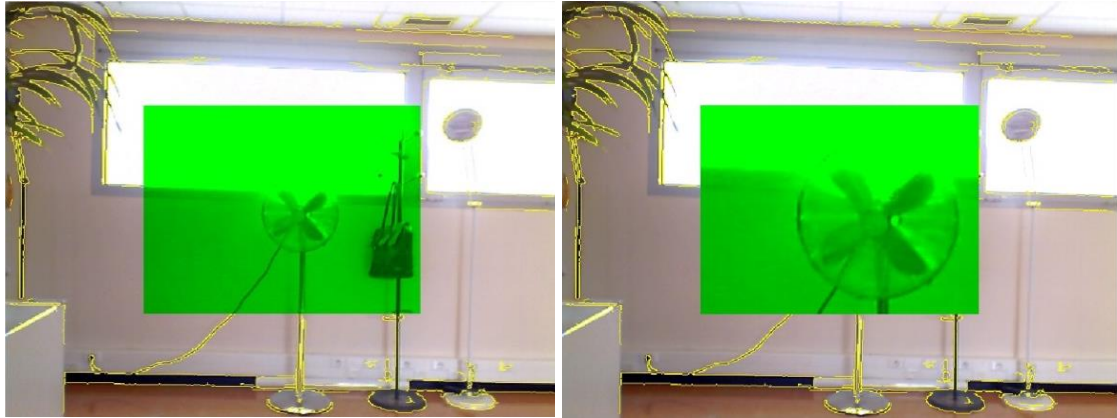


Figure 5.25 : Loupe centrée en vert : (à gauche) grossissement de 1 ; (à droite) grossissement de 2.

5.3.2. Comparaison des systèmes

Les étapes précédemment décrites présentent un intérêt en termes de coût calculatoire. En effet, comme la taille de la zone utile dans l'image (partie de l'image restituée par le module optique) est petite par rapport à la taille de l'image à afficher (image à afficher sur l'écran du dispositif), la détection des contours a donc un coût moins élevé. Par contre, cela a pour effet d'élargir les contours à l'affichage, et de fait, les traits sont alors moins fins que prévu. J'appelle cette solution « implémentation simplifiée ».

J'ai également développé une version prenant en compte la taille du champ utile dans l'image issue de la caméra. Cette zone utile est ensuite redimensionner à la taille de l'image nécessaire pour l'affichage (image restituée par le module optique). Avec cette solution, on peut également réaliser la détection des contours, la superposition de ceux-ci, puis l'affichage de l'image résultat. Cette solution présente l'avantage de respecter la taille des traits prédéfinis pour l'affichage. Cela nécessite cependant plus de capacités de calcul. J'appelle cette solution « implémentation complète ».

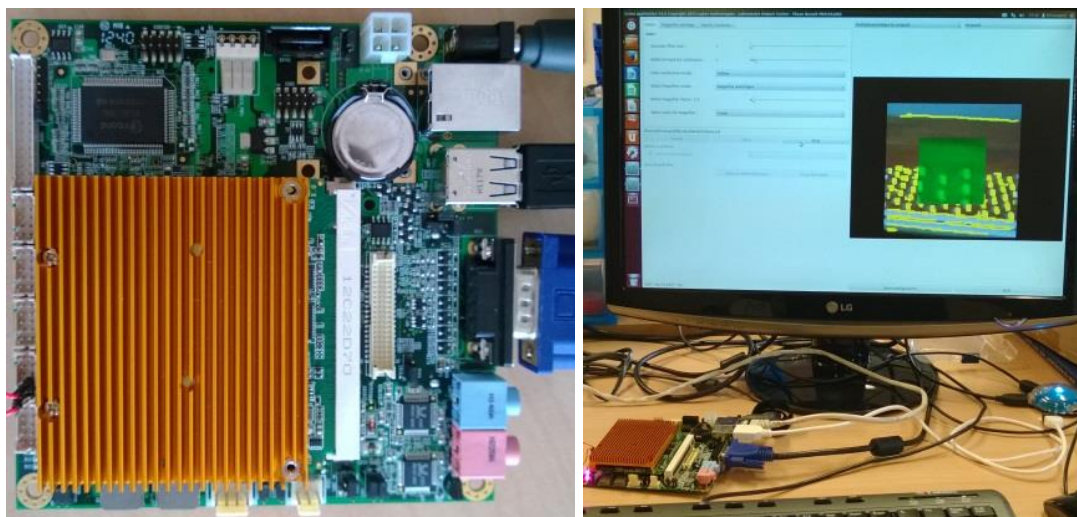


Figure 5.26 : Dispositif embarqué : (à gauche) la carte ; (à droite) le dispositif complet de test (carte + écran + périphériques).

Afin de comparer les temps de calculs de ces deux solutions, je les ai implémentés sur deux dispositifs intégrant des processeurs différents, après avoir fait une étude comparative des différentes solutions et processeurs possibles et adaptés à mon application :

- Un ordinateur portable : DELL VOSTRO comprenant 6Go DDR3 RAM, un processeur Intel Core i5 3210M 2 x 2,50GHz, une carte graphique Nvidia GeForce GT 630 M, un disque dur de 500Go à 5400tr/min et Windows 7 professionnel 64 Bit
- Une carte embarquée : format QSeven équipé de 1 Go DDR3, d'un processeur AMD Fusion G-T40E (APU), une carte SD de 64Go classe 6, cf. Figure 5.26.

Pour comparer les temps de traitement sur ces architectures, j'ai utilisé une même vidéo : résolution 640 X 480 pixels en couleurs à 30 FPS. Cette vidéo représente des déplacements dans une cour, un escalier, un couloir. J'ai utilisé le même code C++ des algorithmes et la librairie de traitement d'images OpenCV⁶⁵. Ce code a été compilé pour le Linux Ubuntu 12.04.2 LTS pour chaque plateforme, en autorisant l'accès des différents cœurs des processeurs avec une librairie générique pour le parallélisme d'application OpemMP⁶⁶. J'ai mesuré les temps de traitement de l'application sur chacune de ces architectures, cf. Figure 5.27 pour la version simplifiée et Figure 5.28 pour la version complète.

⁶⁵ <http://opencv.org/>

⁶⁶ <http://openmp.org/wp/>

	PC portable	Carte embarquée
MIN (en seconde)	0.01	0.04
MAX (en seconde)	0.03	0.10
MOYENNE (en seconde)	0.01	0.06
FPS	100.00	16.80

Figure 5.27 : Comparaison des temps de traitement en fonction de la plateforme de calcul pour « l'implémentation simplifiée ».

	PC portable	Carte embarquée
MIN (en seconde)	0.03	0.21
MAX (en seconde)	0.11	0.35
MOYENNE (en seconde)	0.05	0.24
FPS	20.49	4.12

Figure 5.28 : Comparaison des temps de traitement en fonction de la plateforme de calcul pour « l'implémentation complète ».

Pour les deux implémentations, la partie détection de contours est la plus couteuse en terme de capacité de calcul, suivi par la labellisation et la génération de la carte de contours finale.

Il est clair qu'il y a un facteur proche de 5 entre les temps de traitement entre l'ordinateur portable et la carte embarquée. J'ai également pu vérifier qu'en fonction du compromis performance-qualité des contours, l'outil peut fonctionner à plus de 15 images par seconde.

Lors de ce test de performance, j'ai seulement utilisé la partie « CPU » des processeurs, je n'ai pas cherché pour le moment à bénéficier des accélérations possibles via les calculs massivement parallèles sur le module 3D. Ce qui est néanmoins possible sur l'ordinateur portable en utilisant la carte graphique afin de réaliser plus rapidement certains de mes traitements. De même, il est possible sur le dispositif embarqué d'utiliser ce type d'implémentation pour accélérer les traitements. Notons enfin que la gestion de l'interface graphique consomme un peu de puissance de calcul, ce qui ajoute environ 10ms au temps de calcul.

Il existe donc un certain nombre de voies d'optimisation des temps de traitement qui in fine pourrait me permettre d'atteindre 30 images par seconde sur la carte embarquée : en optimisant la gestion de l'interface graphique, en optimisant l'implémentation de la partie traitement d'image et en sollicitant les accélérations matérielles du module 3D.

5.4. Conclusion

Comme je viens de le présenter, la conception d'un dispositif de réalité augmentée pour l'assistance aux malvoyants est complexe. Il faut à la fois prendre en compte le dispositif d'affichage, les capteurs, et le système de traitement des informations.

Il existe un grand nombre de solutions d'affichage : réalité virtuelle, réalité augmentée, dispositif monoculaire, dispositif binoculaire. Cependant, toutes ces solutions ne sont pas adaptées pour la mobilité des personnes malvoyantes. Au vu de l'état de l'art et des données présentées dans les chapitres 2, 3 et 4, il est à la fois important que le dispositif soit à verres transparents pour permettre la vision de la réalité, puisse disposer d'un grand champ de vision et d'un grand champ d'affichage. Nous avons également vu que pour un certain nombre de malvoyants, un dispositif d'affichage binoculaire est préférable, pour plus de confort visuel et pour l'absence de rivalité binoculaire (phénomène de gêne pouvant apparaître lorsqu'un seul œil observe l'image enrichie). Ce type de dispositif peut être facilement adapté en dispositif monoculaire pour les personnes ne pouvant pas utiliser l'affichage binoculaire.

De même, il existe un certain nombre de technologies de caméras. Afin d'obtenir une image stable, sans déformation due aux mouvements de la caméra et artefact ou saturation du capteur, il est nécessaire de choisir un capteur adapté. En conséquence, afin de pouvoir utiliser à la fois la caméra en intérieur et extérieur, j'ai choisi d'utiliser un capteur CMOS « global shutter » comme le capteur Aptina MT9V024.

Enfin, nous avons vu que pour une application mobile de réalité augmentée, il faut trouver un juste compromis entre la performance de l'algorithme, le poids du calculateur et son autonomie. Pour la mobilité des malvoyants, comme nous l'avons vu au chapitre 1, le poids et la taille du calculateur et des lunettes d'affichage sont trois paramètres primordiaux. Par le biais des expérimentations décrites dans le chapitre 5, nous savons que notre outil de traitement d'images peut être déployé et peut fonctionner sur un ordinateur portable ou sur un système de calcul miniature. Ce deuxième dispositif présente l'avantage d'être intégrable dans un boîtier portable léger pour une autonomie de quelques heures, ce qui dans le cadre de notre application correspond au besoin.

L'objectif à court terme sera d'intégrer le dispositif d'affichage binoculaire, une caméra CMOS Global shutter et de finaliser le boîtier de calcul. Une fois ce dispositif réalisé, je devrai réaliser, de nouveau, une campagne d'évaluation en mobilité afin, de valider l'apport de ce type d'aide pour la mobilité des malvoyants.

Synthèse

Les travaux de ma thèse présentée ici permettent d'apporter les réponses suivantes :

- Définir les signes visuels importants pour les malvoyants à rehausser :
 - Quels sont les indices visuels les plus utiles aux malvoyants ? Les personnes malvoyantes utilisent différents indices visuels dont les contours afin de distinguer les différents éléments situés à proximité d'eux.
 - Existe-t-il un lien entre leurs pathologies et cette réponse ? Il existe un lien entre les besoins et les pathologies. Certaines pathologies comme la DMLA privent le patient de la vision des détails.
 - Quels sont leurs usages des aides visuelles ? Certaines aides sont utilisées pour un guidage lors de tâche en mobilité, les autres aides sont surtout utilisées pour la lecture et l'écriture.
 - Ces aides visuelles répondent-elles à leurs besoins ? Ces aides répondent aux besoins en vision de près, mais les malvoyants rencontrés expriment des besoins sans solutions acceptables pour leur mobilité.
- Évaluer les méthodes existantes pour détecter ces indices visuels :
 - Quel est l'apport du rehaussement des contours pour les malvoyants ? Cet apport est important : les contours peuvent être utilisés comme méthode d'assistance visuelle pour les malvoyants, 80 % des malvoyants ont au moins une fois exprimé obtenir un apport par les contours.
 - Quel est l'apport de ce rehaussement avec les méthodes existantes ? L'apport de ce rehaussement lorsqu'il est réalisé par la méthode Sobel est de 61 %, pour le détecteur de Canny de 47 % et pour le Canny Couleur de 39 %.
- Développer et valider une nouvelle méthode de détecter si nécessaire :
 - Comment réaliser un détecteur de contours adapté aux besoins des malvoyants ? Il est possible en utilisant une pyramide de concernant d'une part que les contours les plus présents dans l'image et d'autre part de fournir un curseur permettant de sélectionner la quantité de contours à restituer en temps réel.

- Quel est l'apport du rehaussement dans ce cas ? L'apport de cette méthode est de 65 % par rapport aux images sources et de 87 % par rapport à Sobel et 96 % par rapport au détecteur de Canny.
- Définir les moyens de rehausser ces indices visuels pour apporter une aide aux malvoyants :
 - Comment (taille, couleurs, forme) doit-on rehausser les contours ? L'ensemble des paramètres de restitution (couleurs, profil et épaisseur) semble être utilisé par les malvoyants pour afficher les contours sur les images originales. Certains malvoyants privilégient le jaune et le blanc principalement. Ces paramètres pourront être adaptés par le malvoyant.
 - Quel lien existe-t-il entre le rehaussement à faire et les différentes atteintes de la vision (champ visuel, acuité, ...) ? Oui, il existe des liens entre les paramètres de rehaussement et les pathologies. Les malvoyants, atteints de champ de vision tubulaire, privilégient le jaune et le blanc principalement. De même, l'épaisseur des traits est adaptée par les malvoyants en fonction des images, de la luminosité, de leur acuité et de leur champ visuel.
- Identifier les composants et matériels nécessaires pour la réalisation de l'aide visuelle finale :
 - Quel type de dispositif d'affichage utiliser ? Il est important que le dispositif soit à verres transparents pour permettre la vision de la réalité, de disposer d'un grand champ de vision et d'un grand champ d'affichage. Un dispositif d'affichage binoculaire est préférable, pour plus de confort visuel et pour l'absence de rivalité binoculaire.
 - Quel type de caméra utiliser ? Il est d'utiliser un capteur CMOS « global shutter » comme le capteur Aptina MT9V024.
 - Quel type de système de calcul permet de répondre à ce besoin ? Il faut trouver un juste compromis entre la performance de l'algorithme, le poids du calculateur et son autonomie. Par le biais des expérimentations, nous savons que les outils proposés peuvent être embarqués sur un système au format QSeven équipé de 1 Go DDR3, d'un processeur AMD Fusion G-T40E (APU), intégrable dans un boîtier portable léger pour une autonomie de quelques heures, ce qui dans le cadre de notre application correspond au besoin.

Conclusion

Dans le premier chapitre de cette thèse, j'ai introduit les notions relatives au système visuel humain, utiles à la compréhension des parties suivantes.

Dans le deuxième chapitre, grâce à une étude approfondie, j'ai recensé les besoins des malvoyants afin de déterminer les informations, les zones, et les contenus les plus pertinents pour réaliser des tâches visuelles, notamment en vision de près, mais aussi en vision intermédiaire et en vision de loin. Lors de cette phase, j'ai fait une immersion en milieu clinique (au sein de l'ARAMAV) afin d'observer le comportement des malvoyants, de déterminer comment ils perçoivent, interagissent, avec les données environnantes quand ils accomplissent une tâche visuelle. J'ai réalisé des entretiens avec des malvoyants et des professionnels du domaine. J'ai également réalisé une analyse des dispositifs du marché et orienté mes entretiens en fonction de ces dispositifs. J'ai notamment pu me rendre compte à quel point une immersion en environnement clinique dédié à la réadaptation visuelle était primordial pour pouvoir bien comprendre quelles sont les techniques apprises par les malvoyants pour compenser leur déficience visuelle, comment prendre en compte ces techniques afin de concevoir une aide d'assistance visuelle basée réalité augmentée qui réponde aux besoins des malvoyants. Dans cette étude, j'ai mis en évidence les nombreux besoins des malvoyants, et démontré que les solutions actuelles présentes sur le marché ne répondent pas à ces besoins, notamment en matière d'aide à la mobilité. Alors qu'il existe actuellement des aides visuelles adaptées à la lecture et à la vision de près de manière générale, il n'existe pas de dispositif approprié à la mobilité. J'ai en particulier démontré qu'il était primordial dans ce contexte d'utiliser les capacités visuelles résiduelles du malvoyant et ses facultés à rechercher par lui-même les informations visuelles dont il a besoin quelle que soit la tâche visuelle et la situation dans laquelle il se trouve. Il est, en effet, primordial que le malvoyant utilise activement ses capacités visuelles résiduelles, de manière autonome, et puisse interagir avec le dispositif d'aide qu'il utilise afin de pouvoir réaliser différents types de tâches visuelles.

Dans les chapitres 3 et 4, j'ai mis en évidence que l'information de contour constitue une information essentielle pour les malvoyants, quelle que soit atteinte de leur champ visuel. J'ai démontré que cette information peut être utilisée pour enrichir leur vision. J'ai également présenté un algorithme de détection de contours qui répond à ce besoin. Pour démontrer la

pertinence de ce détecteur, j'ai d'abord effectué un état de l'art, puis j'ai effectué une étude approfondie des méthodes d'extractions de contours existantes, et enfin analysé et comparé ces détecteurs de contours afin de retenir la méthode la plus appropriée par rapport au cadre applicatif visé. J'ai étudié quel(s) type(s) de contours correspondent le mieux aux besoins des malvoyants et quel(s) critère(s) d'étude décrivent au mieux la « qualité » de ces contours. Dans cette étude, j'ai mis en évidence que la « qualité » d'un contour dépendait de plusieurs paramètres et que ces paramètres n'étaient pas les mêmes pour une personne ayant une vision standard et pour un malvoyant. J'ai aussi démontré que cela n'avait pas de sens de définir un ou plusieurs critères de « qualité » de contours, car d'une tâche visuelle à l'autre, d'une situation à l'autre, les besoins des malvoyants varient. D'où la nécessité de concevoir un détecteur de contours adaptatif, paramétrable, ajustable par l'observateur. La méthode que j'ai implémentée répond à ce besoin, elle est flexible, elle peut s'adapter et être adaptée à la situation et/ou en fonction des besoins du patient. Les expérimentations et tests que j'ai réalisés avec des malvoyants démontrent l'adéquation de ce détecteur de contours par rapport aux besoins des malvoyants et la performance de ce détecteur par rapport à d'autres détecteurs de l'état de l'art. Malgré l'absence de données de référence, de vérité terrain dans le domaine, j'ai réalisé une analyse qualitative qui m'a permis de valider l'hypothèse selon laquelle les contours constituaient une donnée essentielle pour les malvoyants et l'hypothèse selon laquelle la méthode de détection de contours développée est plus pertinente que toutes les autres méthodes en matière d'adaptation aux besoins des malvoyants. Cette méthode est plus pertinente car elle est paramétrable (la quantité de contours, leur épaisseur, leur couleur est ajustable) et donc adaptée aux besoins exprimés.

Dans le cadre d'étude du chapitre 4, j'ai limité mon étude aux images fixes car malgré le fait que la méthode de détection de contours proposée puisse être étendue aux vidéos (cf. chapitre 5) je n'ai pas eu la possibilité d'effectuer des expérimentations en mobilité avec des malvoyants pour tester la robustesse et la performance de ma celle-ci. En effet, pour réaliser ce type d'expérimentation en mobilité il faut d'une part concevoir un protocole d'étude particulièrement complexe (ce qui a été fait) et d'autre part faire valider ce protocole par le corps médical (ce qui est en cours, mais prend un certain temps). Il est à noter que je me suis déjà confronté à ce type de problème (de délais) lors de mes premières expérimentations (à partir d'images fixes affichées sur écran CRT ou LCD) décrites dans le chapitre 4. Cette évaluation en mobilité sera à réaliser après la thèse.

Dans le cinquième chapitre, j'ai démontré que les dispositifs de réalité augmentée étaient plus appropriés à la problématique visée que les dispositifs de réalité virtuelle. J'ai mis en évidence les principaux critères techniques à prendre en compte pour concevoir et développer un système d'aide visuelle adapté aux besoins des malvoyants. Je me suis notamment attaché à implémenter une solution logicielle qui permet d'afficher en temps réel des informations image sur un dispositif de réalité augmentée. Je me suis également attaché à développer un générateur d'images destiné à afficher les contours, de manière confortable, sur un dispositif binoculaire. Au-delà des problèmes de recalage d'images, de compensation des distorsions géométriques et de prise en compte des données physiologiques propres à chaque observateur, je me suis également attaché à démontrer que la méthode de détection de contours que je propose pouvait fonctionner en temps réel sur un dispositif mobile autonome. J'ai également implémenté une interface afin que les malvoyants puissent interagir avec le dispositif. Dans le chapitre 5, j'ai également suggéré quelques pistes pour optimiser les temps de calcul de l'algorithme de détection de contours et améliorer la performance de la méthode utilisée en exploitant par exemple la dimension temporelle. Il semble utile d'exploiter l'information de profondeur (un malvoyant étant, par exemple, plus intéressé à localiser une porte dans un couloir quand celle-ci est proche de lui plutôt qu'une autre plus éloignée), mais cette hypothèse mérite d'être vérifiée auprès des malvoyants.

Concernant les aspects optimisation des temps de traitement (sur système embarqué) et amélioration de la méthode de détection (via une technique de type SLAM), j'ai contribué activement au dépôt d'un projet ANR, dans le cadre de l'appel à projet ANR TecSan 2013, ce projet n'a pas été retenu mais a été favorablement jugé par les experts qui l'ont analysé. Ce projet a été légèrement remanié, puis re-soumis lors du dernier appel à projet de l'ANR.

Concernant les publications scientifiques, compte tenu des contributions et avancées relatives à ce travail de thèse, certains aspects jugés confidentiels par Laster Technologies ne pourront être valorisés directement sous forme de publications, néanmoins d'autres aspects ont pu donner lieu à publication, cf. article soumis à Transactions on Accessible Computing⁶⁷ (ACM Taccess) en décembre 2013 et article accepté à HCI 2014⁶⁸.

Pour finir, j'ai également contribué en 2013 au dépôt d'un dossier soumis à l'appel à projet « CNRS-Oxford scheme », cette proposition a été acceptée. Via le financement obtenu, je vais pouvoir effectuer un séjour de 1 mois à Oxford dans le Nuffield Department of Clinical

⁶⁷ <http://www.rit.edu/gccis/taccess/>

⁶⁸ <http://www.hcii2014.org/>

Neuroscience afin d'une part d'évaluer les performances de mes algorithmes avec des malvoyants dans un contexte de mobilité et d'autre part comparer mes travaux avec ceux de l'équipe de Stephen Hicks.

À court terme, je dois réaliser les essais (en environnement clinique) du dispositif de réalité augmentée qui a été développé, afin de valider la pertinence et l'adéquation aux besoins de ce dispositif dans le contexte de la mobilité. Cet essai en mobilité est en train d'être mis en place et a démarré en février 2014. Les premiers résultats de cette étude devraient être connus au printemps 2014. Ces résultats permettront de quantifier l'apport des contours et du détecteur présenté dans ce document pour l'assistance visuelle des malvoyants en utilisant une paire de lunettes de réalité augmentée. Une fois que cette validation en mobilité aura été effectuée, il sera alors possible de compléter le nombre d'informations utiles pour le malvoyant, telles que l'information de profondeur, à afficher sur le dispositif de réalité augmentée. Je pourrai développer d'autres fonctionnalités (permettant de répondre à d'autres besoins des malvoyants) pour lesquelles il n'existe actuellement pas de solution. L'évolution des technologies permettant d'augmenter l'autonomie de ce type de dispositifs, tout en réduisant sa taille, devrait encore contribuer à améliorer significativement l'apport de cette solution pour les malvoyants.

Enfin, mes travaux pourront être utilisés dans d'autres domaines que celui de l'assistance visuelle des malvoyants. Les méthodes d'analyse des besoins et de validation d'une solution par rapport à ces besoins que j'ai mis en œuvre, dans un contexte de développement de solution d'assistance à base d'un dispositif de réalité augmentée, pourraient être transposées à d'autres contextes applicatifs, tel que l'assistance à la personne, l'industrie, la maintenance et le secteur militaire.

On pourrait notamment utiliser le système de calcul pour réaliser beaucoup de traitements d'images embarquées, soit dans des produits de réalité augmentée afin de permettre le fonctionnement de chaque application en autonomie et en mobilité. De plus, ce type de développement peut également être utilisé pour de la réalité virtuelle. Une unité de calcul portable peut être utilisée pour résoudre un grand nombre d'applications nécessitant de pouvoir réaliser une analyse de données, d'images en temps réel et/ou en mobilité afin d'améliorer les conditions de travail de la personne ou de lui apporter des informations nécessaires supplémentaires.

Le dispositif d'affichage binoculaire que j'ai développé pour ce dispositif d'assistance peut directement être utilisé pour beaucoup d'applications à destination des professionnels, ou pour

des dispositifs d'affichage pour le grand public. Il permet, en effet, d'afficher des informations en binoculaire, en utilisant les deux yeux de la personne, des informations en réalité augmentée. Ce type d'affichage est moins perturbant que l'affichage sur un seul œil des informations ; cela permettra une plus grande acceptation des dispositifs de réalité augmentée. Ce générateur d'images peut être paramétré en fonction du module optique utilisé pour afficher les images pour les yeux gauches et droits. Il est donc facilement utilisable avec tous les modules optiques de l'entreprise Laster Technologies pour créer une ou des versions binoculaires de chaque module optique. Enfin, il est possible, en utilisant ce générateur d'image, d'adapter le taux de recouvrement des images gauches et droites afin de créer une image enrichie plus ou moins grande en fonction de l'application visée. Cette adaptation nécessite de conserver tout de même un taux de recouvrement suffisant afin de maintenir la fusion des images par le système visuel humain. En respectant la fusion des images, il est ainsi possible de créer plusieurs versions d'un même produit binoculaire proposant un champ d'affichage plus ou moins étendu en fonction de l'usage souhaité.

Je pourrai également utiliser l'algorithme de détection de contours (celui-ci étant paramétrable) pour réaliser d'autres solutions pour malvoyants. En effet, le rehaussement des contours peut également être utile au sein d'autres dispositifs d'aide tels que les logiciels agrandisseurs, ou les loupes électroniques afin de ne pas proposer uniquement un agrandissement des images, mais des rehaussements de celles-ci. Ce détecteur de contours a été développé pour les malvoyants mais pourrait très bien être utilisé par les militaires, les pilotes d'avion, les pompiers, les conducteurs, car il permet de détecter des contours en toute situation. Dans les environnements où la visibilité est réduite telle que le brouillard, de la fumée, la nuit, l'apport de contours en réalité augmentée peut permettre d'améliorer la perception de la situation et ainsi faciliter la prise de décision.

De plus, ce détecteur de contours pourra être utilisé dès que l'on ne connaît pas nécessairement l'échelle d'information requise ou que l'application requiert une adaptation des paramètres du détecteur en fonction de l'objet, de la situation, de l'image. Je peux notamment envisager des applications en vision industrielle pour le contrôle qualité de produits comprenant des textures très diverses sur une même ligne de production telle que l'inspection de meuble en bois, d'objets décoratifs...

Je peux également envisager des usages de ce détecteur pour adapter et améliorer des outils de suivi d'objets, de calcul des points d'intérêts de calculs de flux optique. La connaissance des contours les plus importants pour le système visuel dans une image ou dans une scène peut

permettre de ne retenir que ces éléments lors de la détection et le suivi d'objets définis. De même la prise en compte uniquement des éléments les plus pertinents au niveau du système visuel peut permettre de réduire la quantité d'informations à détecter et à analyser. L'usage des éléments les plus caractéristiques, uniquement, pourrait permettre d'améliorer les performances et les temps de calculs d'algorithmes de détection d'objets, de reconnaissance d'objets, de calcul de flux optique.

Annexe 1 : Questionnaire

Partie A : identité :

Vous êtes : une femme, un homme

Votre âge :

Combien de personnes vivent avec vous (vous excepté) : _____

Vous êtes : travailleur, au chômage, retraité, en invalidité

autre : _____

Depuis combien de temps êtes-vous malvoyants ? _____

Partie B : Livres :

- Connaissez-vous les livres agrandis ? Oui, Non
- Utilisez-vous les livres agrandis ? Oui, Non
- Connaissez les lieux pour acheter des livres agrandis ? Oui, Non
- Connaissez-vous les livres audio ? Oui, Non
- Utilisez-vous les livres audio ? Oui, Non
- Connaissez les lieux pour acheter des livres audio ? Oui, Non

Partie C : utilisation d'aides techniques :

- Les aides vocales (reconnaissance, synthèse) :
 - Utilisez-vous des aides ? Oui, Non
 - Savez-vous comment utiliser ces aides ? Oui, Non
 - Connaissez-vous les lieux pour acheter de telles aides ? Oui, Non
 - Êtes-vous intéressé pour acquérir ce type d'aides ? Oui, Non
 - Utilisez-vous ces aides ? Rarement ou jamais, quelques fois, souvent, tous les jours
- Les aides optiques (loupes, système-télescopique, ...) :
 - Utilisez-vous des aides ? Oui, Non
 - Savez-vous comment utiliser ces aides ? Oui, Non

- Connaissez-vous les lieux pour acheter de telles aides ? Oui, Non
- Êtes-vous intéressé pour acquérir ce type d'aides ? Oui, Non
- Utilisez-vous ces aides ? Rarement ou jamais, quelques fois, souvent, tous les jours
- Les loupes électroniques :
 - Utilisez-vous des aides ? Oui, Non
 - Savez-vous comment utiliser ces aides ? Oui, Non
 - Connaissez-vous les lieux pour acheter de telles aides ? Oui, Non
 - Êtes-vous intéressé pour acquérir ce type d'aides ? Oui, Non
 - Utilisez-vous ces aides ? Rarement ou jamais, quelques fois, souvent, tous les jours
- Les télé-agrandisseurs :
 - Utilisez-vous des aides ? Oui, Non
 - Savez-vous comment utiliser ces aides ? Oui, Non
 - Connaissez-vous les lieux pour acheter de telles aides ? Oui, Non
 - Êtes-vous intéressé pour acquérir ce type d'aides ? Oui, Non
 - Utilisez-vous ces aides ? Rarement ou jamais, quelques fois, souvent, tous les jours
- Aides vocales (GPS,...)
 - Utilisez-vous des aides vocales ? Oui, Non
 - Savez-vous comment utiliser ces aides vocales? Oui, Non
 - Connaissez-vous les lieux pour acheter de telles aides vocales? Oui, Non
 - Êtes-vous intéressé pour acquérir ce type d'aides vocales? Oui, Non
 - Utilisez-vous ces aides vocales ? Rarement ou jamais, quelques fois, souvent, tous les jours
- Loupes informatiques (zoomtext, loupe windows, supernova...) :
 - Utilisez-vous des aides ? Oui, Non
 - Savez-vous comment utiliser ces aides ? Oui, Non
 - Connaissez-vous les lieux pour acheter de telles aides ? Oui, Non
 - Êtes-vous intéressé pour acquérir ce type d'aides ? Oui, Non
 - Utilisez-vous ces aides ? Rarement ou jamais, quelques fois, souvent, tous les jours
- Les logiciels vocaux (jaws, synthèse vocal,...) :

- Utilisez-vous des aides ? Oui, Non
- Savez-vous comment utiliser ces aides ? Oui, Non
- Connaissez-vous les lieux pour acheter de telles aides ? Oui, Non
- Êtes-vous intéressé pour acquérir ce type d'aides ? Oui, Non
- Utilisez-vous ces aides ? Rarement ou jamais, quelques fois, souvent, tous les jours
- Les machines à lire :
 - Utilisez-vous des aides ? Oui, Non
 - Savez-vous comment utiliser ces aides ? Oui, Non
 - Connaissez-vous les lieux pour acheter de telles aides ? Oui, Non
 - Êtes-vous intéressé pour acquérir ce type d'aides ? Oui, Non
 - Utilisez-vous ces aides ? Rarement ou jamais, quelques fois, souvent, tous les jours
- Éclairage :
 - Utilisez-vous des aides ? Oui, Non
 - Savez-vous comment utiliser ces aides ? Oui, Non
 - Connaissez-vous les lieux pour acheter de telles aides ? Oui, Non
 - Êtes-vous intéressé pour acquérir ce type d'aides ? Oui, Non
 - Utilisez-vous ces aides ? Rarement ou jamais, quelques fois, souvent, tous les jours
- Outil braille :
 - Utilisez-vous des aides ? Oui, Non
 - Savez-vous comment utiliser ces aides ? Oui, Non
 - Connaissez-vous les lieux pour acheter de telles aides ? Oui, Non
 - Êtes-vous intéressé pour acquérir ce type d'aides ? Oui, Non
 - Utilisez-vous ces aides ? Rarement ou jamais, quelques fois, souvent, tous les jours

Partie D : les impacts :

Avez-vous réalisé des adaptations dans :

	Oui	Non	Non applicable
--	-----	-----	----------------

votre cuisine			
votre chambre			
votre salle de bain			
salle à manger			
Salon			
Escaliers			
activités journalières (faire le ménage, cuisiner,...)			
vos activités sociales (rencontrer des personnes, téléphoner, internet,...)			
vos activités familiales (reconnaissance de personnes)			
vos loisirs (télévision, musique, sport, mobilité)			
votre poste de travail			

Partie E : vos besoins :

Êtes-vous autonome (faire seul) pour :

	Oui	Non
Vous déplacer sur de courts parcours connus		
Vous déplacer sur de courts parcours inconnus		
Vous déplacer sur des distances longues		
Faire vos courses		
Utiliser internet		
Téléphoner		
Cuisiner		
Vous habiller		
Faire le ménage		
Reconnaître des personnes		
Lire		
écrire		

Indiquer les aides ou fonctions que vous avez besoin :

	Oui	Non
Aide à la lecture		
Aide à la vision des couleurs		
Lecture vocale		
Commande vocale de l'appareil		
Aide de vision de près (travail de précision, bricolage)		
Image agrandie		
Isolation du texte et mise en forme		
Affichage des contours (mise en évidence)		
Affichage des contours en vision nocturne		
Mode vision nocturne		
Accès au téléphone		
Accès à internet		
Regarder un film (DVD, télévision)		
Guidage GPS		

Partie F : les caractéristiques :

Il est énuméré ici une liste de caractéristiques pour une éventuelle aide visuelle, pouvez-vous d'indiquer les caractéristiques à rester ?

	Oui	Non
Affichage couleur		
Affichage binoculaire		
Affichage monoculaire		
Réglage de la luminosité		
Possibilité d'occulté le réel (système immersif opaque sans vision du monde réel)		
Lunettes avec fils		
Lunettes sans fils		
Boitier distant (à la ceinture ou sacoche)		
Reconnaissance vocale (commande vocale)		
Système braille (bouton et ou clavier)		
Lecture de texte (synthèse vocale)		
Mode automatique (utilisation simple semi automatisé)		
Mode manuel (utilisation avec tous les réglages)		

Indiquer toutes les valeurs convenables pour vous ?

Autonomie :

- moins de 2 heures
- entre +2 et 4 heures
- entre +4 et 6 heures
- entre +6 et 8 heures
- Plus de +8 heures

Poids des lunettes :

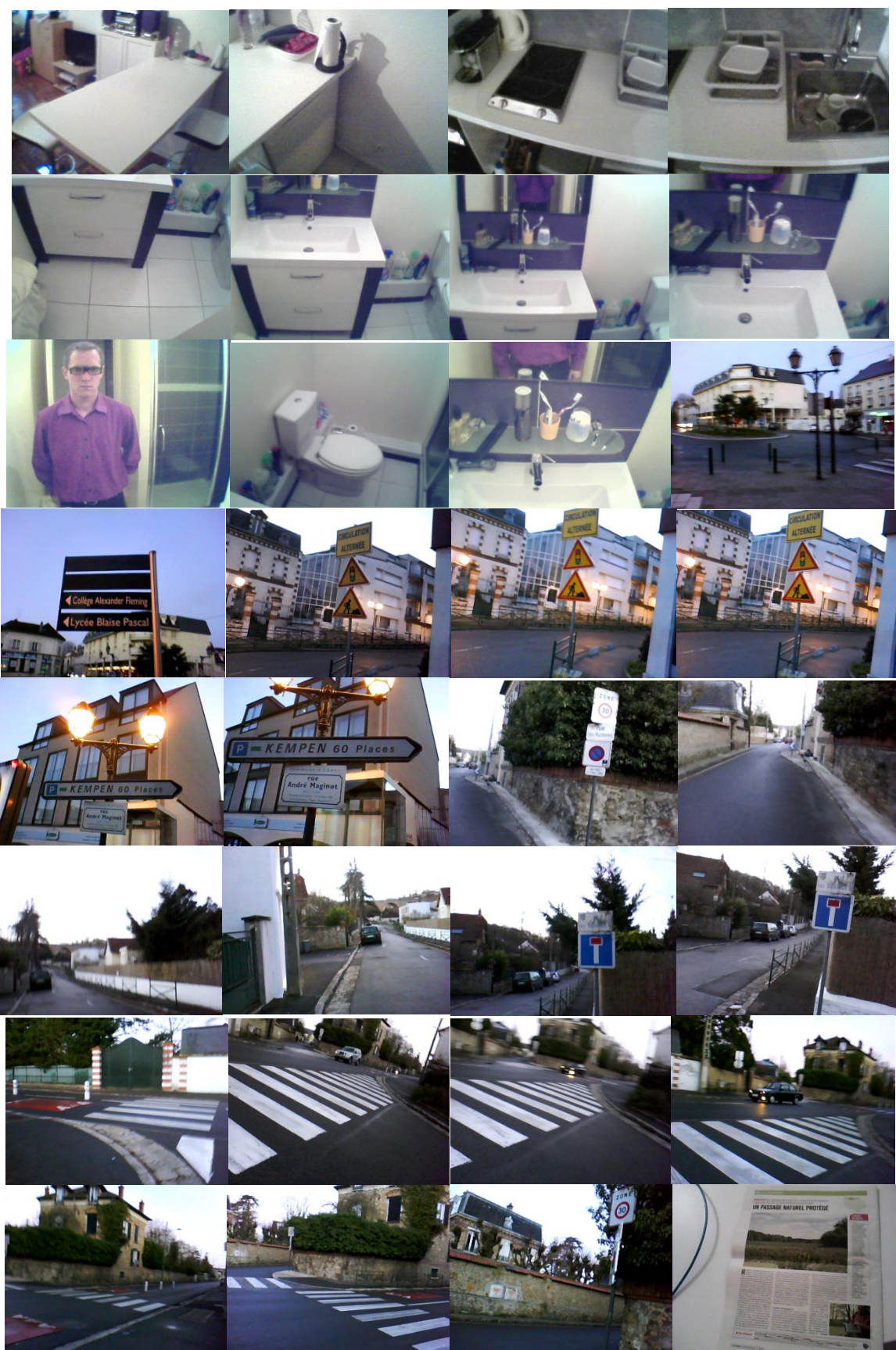
- moins de 99 grammes
- entre 100 et 119 grammes
- entre 120 et 139 grammes
- entre 140 et 159 grammes
- entre 160 et 179 grammes
- Plus de 180 grammes

Poids du boîtier :

- moins de 199 grammes
- entre 200 et 299 grammes
- entre 300 et 399 grammes
- entre 400 et 499 grammes
- entre 500 et 699 grammes
- Plus de 700 grammes

Annexe 2 : la banque d'images





Annexe 3 : Images et vérité terrain

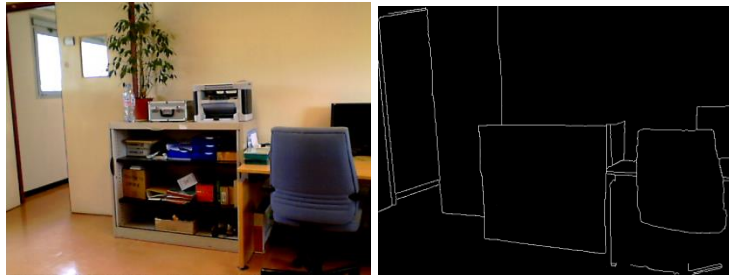


Image 1 et sa vérité terrain



Image 2 et sa vérité terrain

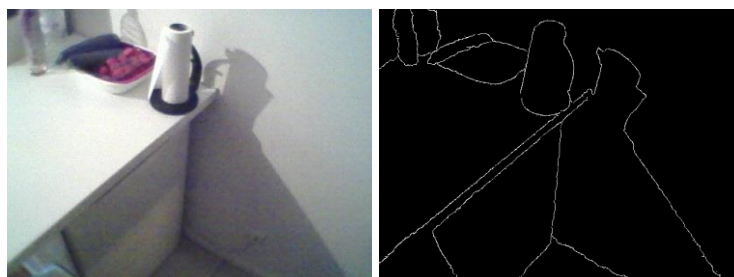


Image 3 et sa vérité terrain



Image 4 et sa vérité terrain

Annexe 4 : Analyse quantitative

J'ai appelé les contours bien détectés TP (True Positive) chaque pixel de notre carte de contours correspondant à un pixel contour dans la carte vérité terrain. Nous avons appelé les contours Faux positifs FP (False Positive) chaque pixel de notre carte de contours ne correspondant pas à un pixel contours dans la carte vérité terrain. Nous avons appelé les contours faux négatifs FN (False Négative) chaque pixel de la carte vérité terrain non détecté. Nous pouvons calculer la proportion de réponses pertinentes parmi toutes les réponses données par le détecteur :

$$\text{Précision} = \frac{TP}{TP + FP}$$

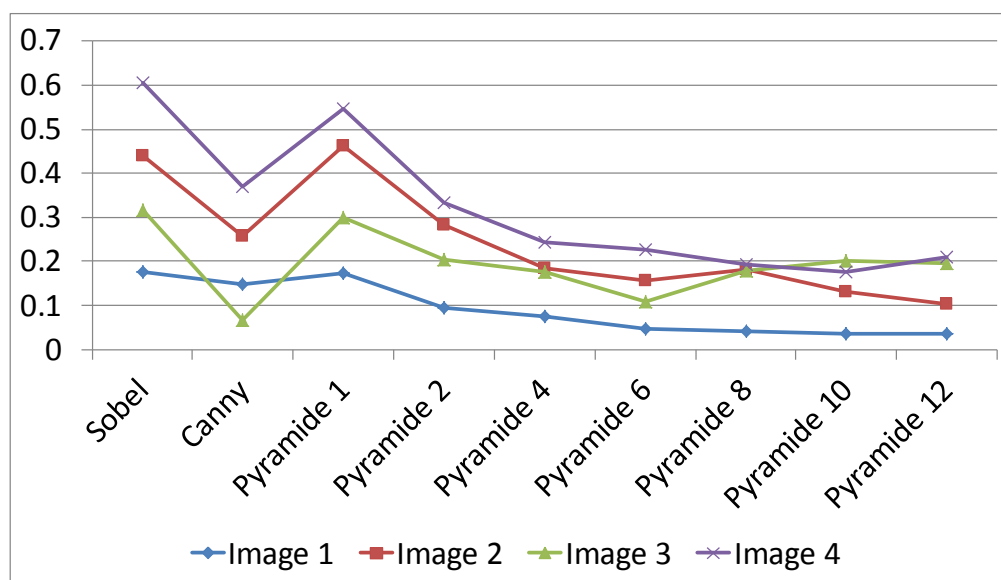


Figure A4.1 : Courbes de l'indicateur Précision pour chaque image en fonction du détecteur de contours.

Comme le montrent les courbes de la Figure A4.1, il existe une différence entre le détecteur de Canny et le filtre de Sobel au niveau de la qualité des contours détectés. Le détecteur de Canny permet de détecter davantage de contours proches de ceux attendus. Les données de la pyramide montrent une variation des valeurs, à quelques exceptions près pour l'image 3, en fonction de l'évolution du paramètre de filtrage.

L'abscisse sur les graphiques indique en premier les méthodes de Sobel et Canny puis, les résultats pour la méthode pyramidale pour différentes valeurs de sigma du filtre gaussien.

Le filtre de Canny a été réglé pour conserver uniquement les contours les plus marqués (les 30% supérieurs de l'histogramme cumulé). Ce réglage peut générer beaucoup de contours. La méthode pyramidale se base également sur ce filtre de Canny avec ces mêmes réglages.

Nous pouvons calculer la capacité du détecteur à ne donner que des réponses pertinentes, cf.

Figure A4.2 :

$$Rappel = \frac{TP}{TP + FN}$$

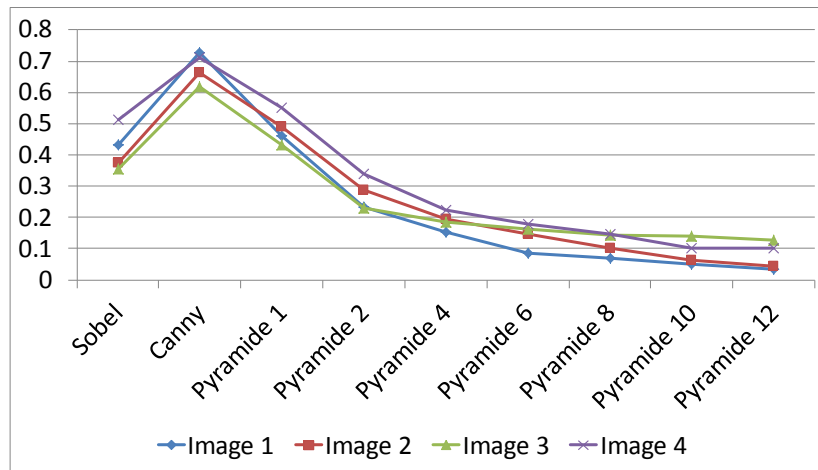


Figure A4.2 : Courbes de l'indicateur Rappel pour chaque image en fonction du détecteur de contours.

Nous pouvons calculer la capacité du système à fournir uniquement des réponses pertinentes en refusant les autres :

$$F - Mesure = \frac{2 * Rappel * Précision}{Rappel + Précision}$$

Comme le montre, les courbes de la Figure A4.3, il existe une différence entre le détecteur de Canny et filtre de Sobel au niveau de la qualité des contours détectés. De même, il est visible que le filtre de Canny détecte beaucoup trop de contours. Les données de la pyramide montrent bien l'évolution de la quantité de contours détectés en fonction du paramètre de filtrage.

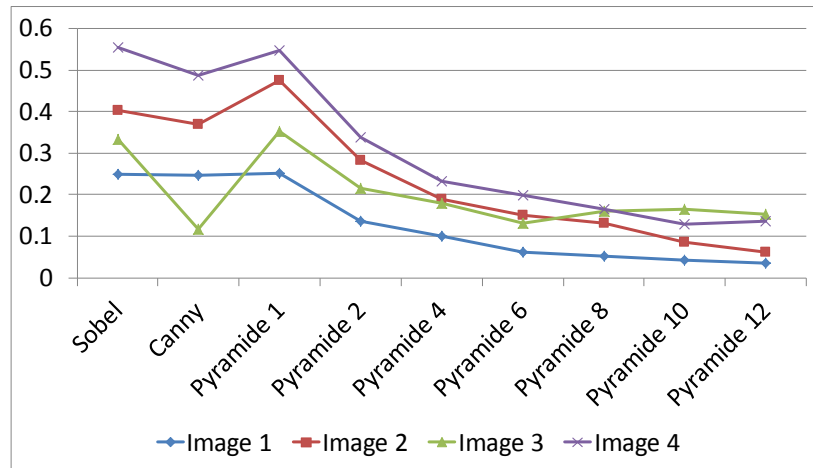


Figure A4.3 : Courbes de l'indicateur F-Mesure pour chaque image en fonction du détecteur de contours.

Comme le montrent, les courbes de la figure A4.3, il existe une différence entre le détecteur de Canny et filtre de Sobel au niveau de la qualité des contours détectés. De même, il est visible que le filtre de Canny détecte les contours souhaités, mais également beaucoup trop de contours. Les données de la pyramide montrent bien l'évolution de la quantité de contours détectés en fonction du paramètre de filtrage.

Nous pouvons calculer la probabilité de pixels non détectés P_{nd} [24] :

$$P_{nd} = \frac{FN}{\text{Max}(NV, NB)}$$

Avec NV le nombre de pixel contours de la vérité terrain et NB le nombre de pixels de notre carte de contours. La valeur idéale est 0. Les valeurs peuvent fluctuer de 0 à 1. Pour chaque image, j'ai pris en compte plusieurs valeurs du paramètre de filtrage de notre méthode.

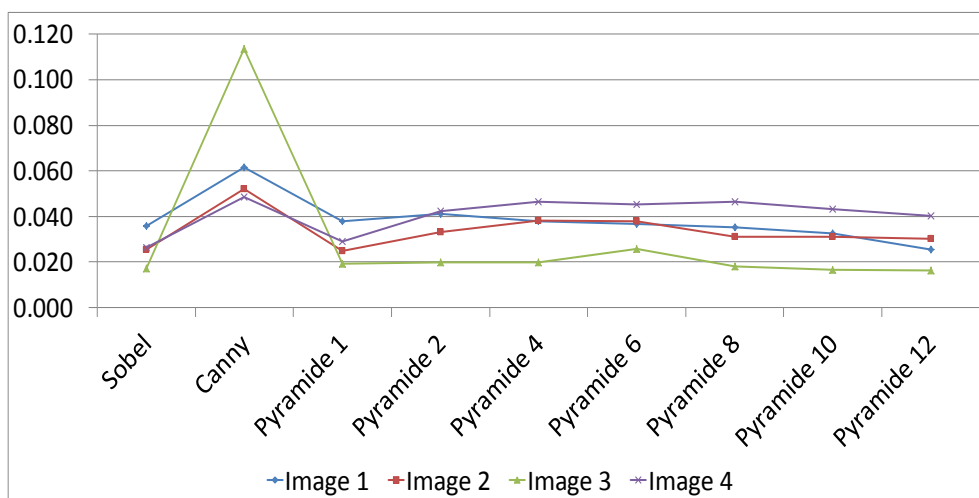


Figure A4.4 : Courbes de l'indicateur Pnd pour chaque image en fonction du détecteur de contours.

D'après les données de la Figure A4.4, le détecteur de Canny détecte beaucoup de contours, la probabilité de ne pas détecter un contour est donc très faible. Cela implique un certain nombre de contours non détectés, ce nombre augmente au fur et à mesure que le paramètre de filtrage augmente. Cela correspond au fonctionnement souhaité pour notre méthode à base de pyramide à savoir la sélection des contours les plus pertinents.

De la même manière, nous pouvons calculer la probabilité de pixels correctement détectés P_{co} [24]:

$$P_{co} = \frac{TP}{Max(NV, NB)}$$

La valeur idéale est 1. Les valeurs peuvent fluctuer de 0 à 1.

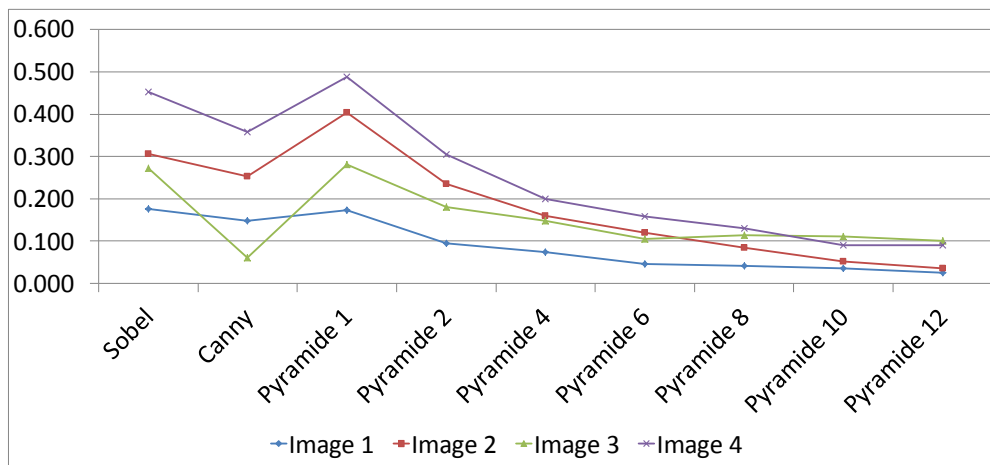


Figure A4.5 : Courbes de l'indicateur P_{co} pour chaque image en fonction du détecteur de contours.

Au vu des valeurs, de la Figure A4.5, pour l'indicateur P_{co} , le détecteur de Canny extrait beaucoup de contours, mais pas forcément les contours idéaux, car sa valeur est plus faible que certaines valeurs obtenues par notre pyramide. La pyramide, comme souhaitée, permet de bénéficier de l'effet de filtrage de la quantité de contours, d'une part et d'autre part de conserver les contours essentiels et correspondants mieux à la vérité terrain. Cette méthode semble plus adaptée pour extraire les contours essentiels dans une scène.

De la même manière, nous pouvons calculer la probabilité de pixels faussement détectés P_{fa} [24], cf. Figure A4.6 :

$$P_{fa} = \frac{FP}{Max(NV, NB)}$$

La valeur idéale est 0. Les valeurs peuvent fluctuer de 0 à 1.

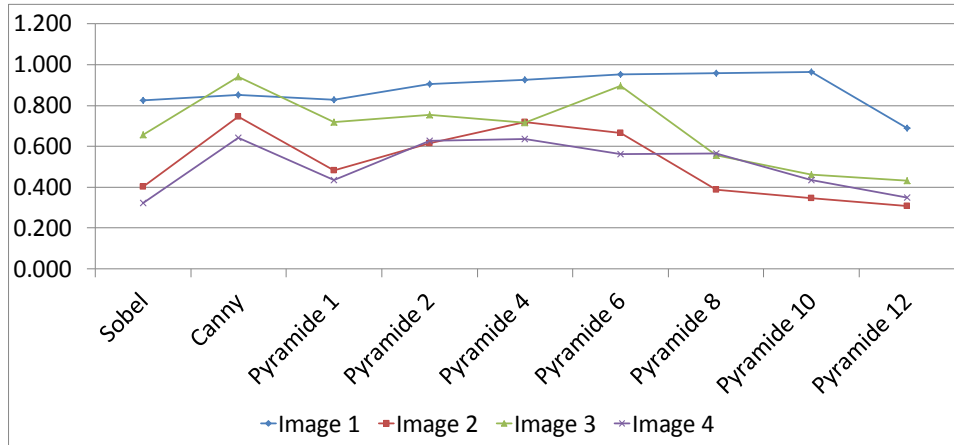


Figure A4.6 : Courbes de l'indicateur Pfa pour chaque image en fonction du détecteur de contours.

Il a également été proposé la mesure du mérite de Pratt [24] :

$$IMP = \frac{1}{\text{Max}(NV, NB)} * \sum_{i=1}^{NB} \frac{1}{1 + \frac{1}{9} * Di^2}$$

Di est la distance en pixel entre un pixel de contours réel et le pixel contour de la vérité terrain le plus proche. Il s'agit de mesurer la distance entre la position du contour idéal à détecter de la vérité terrain et la position du contour détecté. La valeur optimale est 1. Les valeurs peuvent fluctuer de 0 à 1.

D'après les courbes de l'indicateur IMP de la Figure A4.7, la progression de l'indicateur montrant l'éloignement progressif des contours détectés par la pyramide par rapport à la vérité terrain. La pyramide, comme souhaitée, permet de bénéficier de l'effet de filtrage de la quantité de contours, d'une part et d'autre part de conserver les contours essentiels et correspondants mieux à la vérité terrain. Cette méthode semble plus adaptée pour extraire les contours essentiels dans une scène.

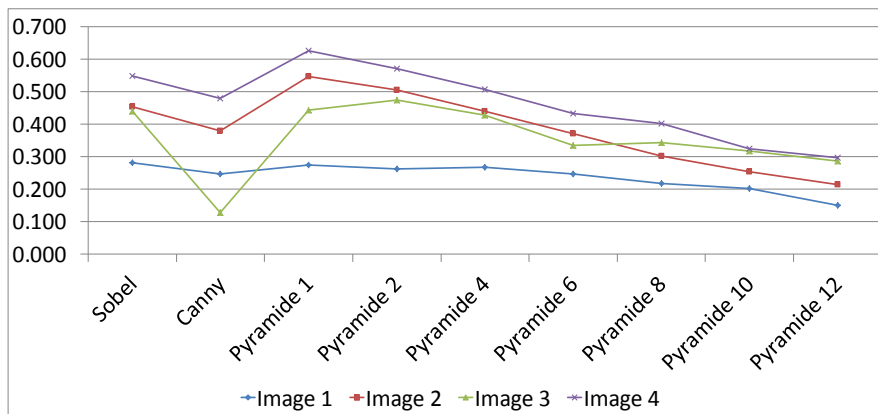


Figure A4.7 : Courbes de l'indicateur IMP pour chaque image en fonction du détecteur de contours.

Il a également été proposé de combiner ces indicateurs dans une seule métrique [24], cf. Figure A4.8 :

$$D4L2 = \sqrt[2]{(Pco - 1)^2 + (IMP - 1)^2 + Pnd^2 + Pfa^2}$$

La valeur optimale est 0, les valeurs peuvent fluctuer de 0 à 2.

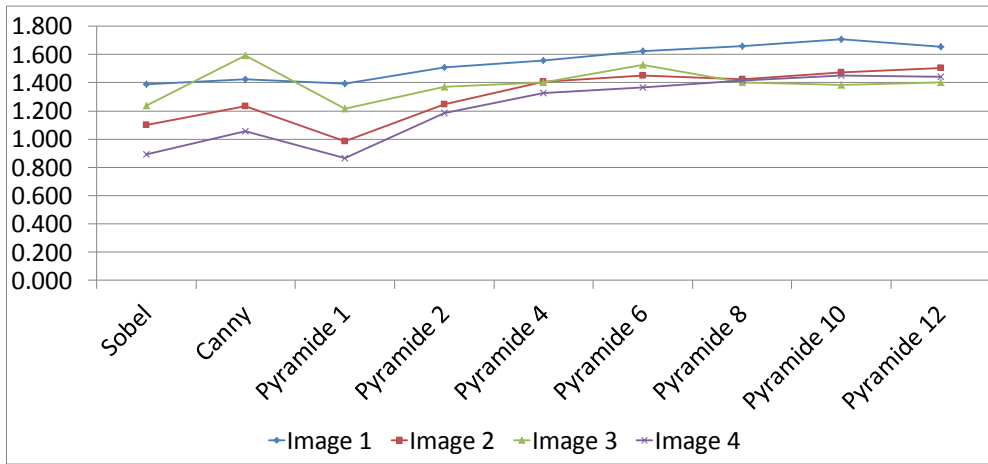


Figure A4.8 : Courbes de l'indicateur D4L2 pour chaque image en fonction du détecteur de contours.

Il a également été proposé la mesure de distance entre deux images [40] :

$$Lx(Carte\ contours/Vérité\ terrain) = \left[\frac{\sum |imagecontours(x) - vérité\ terrain(x)|^q}{Nombre\ de\ pixel\ dans\ l'image} \right]^{1/q}$$

Cela revient à calculer la somme des erreurs élevée à la puissance q, de diviser cette somme par le nombre total de pixels dans l'image et prendre cette valeur à puissance 1/q en d'autre terme cela revient à mesurer l'erreur entre les deux images puis de calculer cette somme à une certaine puissance. La distance l1 correspond à la mesure de l'erreur entre les deux images, la distance L2 correspond à somme quadratique des erreurs entre les deux images. Les distances égales à 0 sont les distances idéales.

Il est nettement visible sur les courbes présentées et en annexe 4, que ces indicateurs sont difficilement exploitables ici, car il nous faut réaliser ces analyses sur une série d'images plus conséquentes, et surtout sur une vérité terrain réalisée auprès d'un plus grand nombre de personnes. Idéalement, la vérité terrain devrait être réalisée avec des personnes malvoyantes. Cela compliquerait l'analyse des résultats, car nous n'aurions pas une mais, des vérités terrain.

Ces indicateurs permettent néanmoins de mettre en évidence un de nos objectifs pour le développement de notre méthode pyramidale à savoir l'adaptation progressive de la quantité

de contours. Il est visible sur un grand nombre de courbes une progression des valeurs indiquant une diminution du nombre de contours détectés, plus le paramètre de filtrage est élevé. La Figure A4.9, la Figure A4.10, la Figure A4.11 et la Figure A4.12 représentent respectivement les valeurs pour les indicateurs L1, L2, L3 et L4.

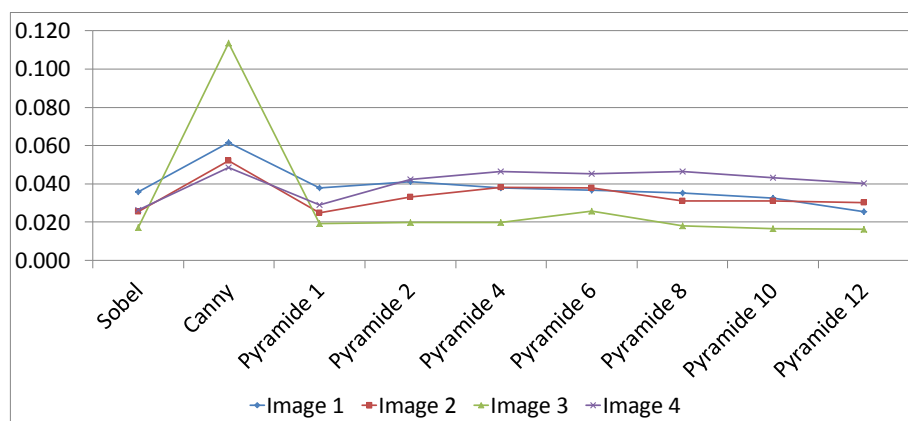


Figure A4.9 : Courbes de l'indicateur L1 pour chaque image en fonction du détecteur de contours.

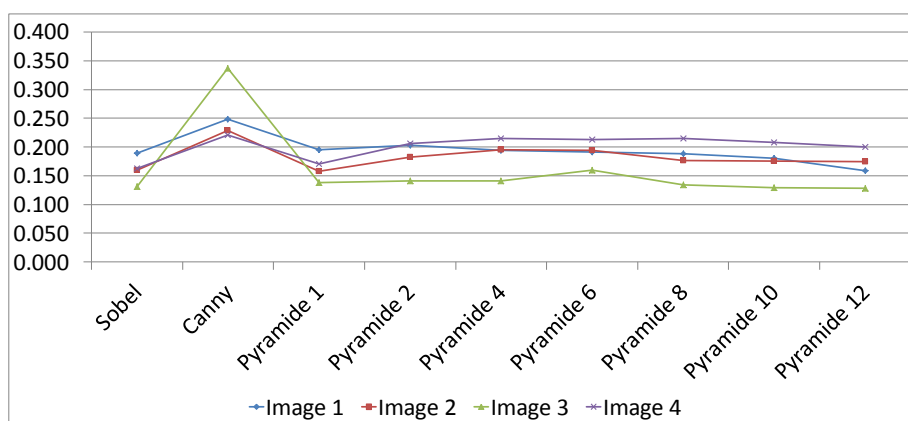


Figure A4.10 : Courbes de l'indicateur L2 pour chaque image en fonction du détecteur de contours.

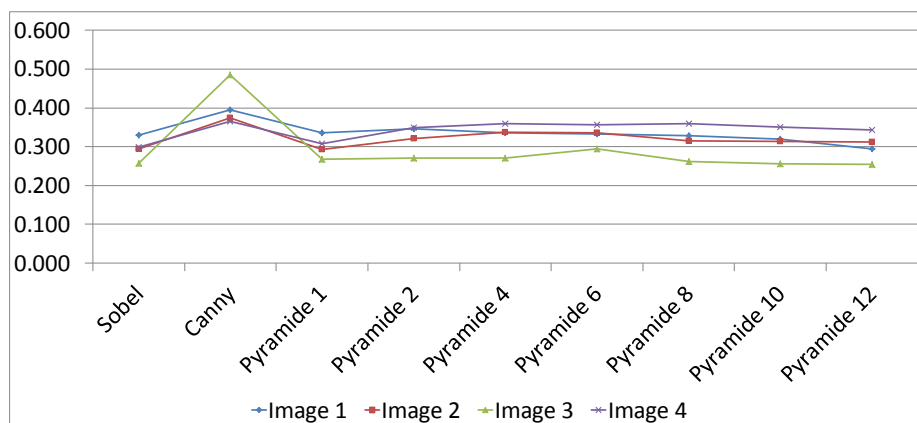


Figure A4.11 : Courbes de l'indicateur L3 pour chaque image en fonction du détecteur de contours.

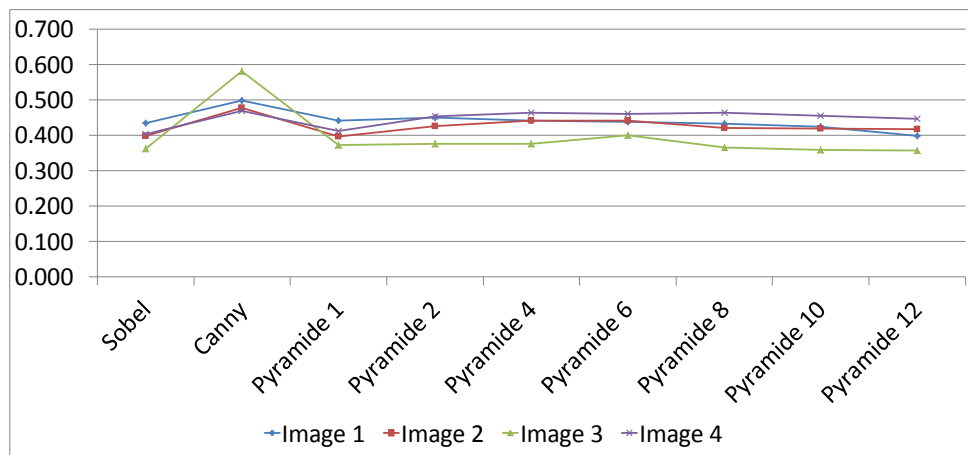


Figure A4.12 : Courbes de l'indicateur L4 pour chaque image en fonction du détecteur de contours.

Bibliographie

- [1] Mc gill university <http://lecerveau.mcgill.ca/>.
- [2] The international classification of diseases, 10th revision (icd-10) 2010. Technical report, Geneva Switzerland, 2010.
- [3] Radhakrishna Achanta, Francisco Estrada, Patricia Wils, and Sabine Süssstrunk. Salient region detection and segmentation. In *Computer Vision Systems*, pages 66–75. Springer, 2008.
- [4] Radhakrishna Achanta, Sheila Hemami, Francisco Estrada, and Sabine Susstrunk. Frequency-tuned salient region detection. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 1597–1604. IEEE, 2009.
- [5] Radhakrishna Achanta and S Susstrunk. Saliency detection using maximum symmetric surround. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pages 2653–2656. IEEE, 2010.
- [6] Carlos Aguilar and Eric Castet. Gaze-contingent simulation of retinopathy: some potential pitfalls and remedies. *Vision research*, 51(9):997–1012, 2011.
- [7] Timothy J. Andrews and David M. Coppel. Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision Research*, 39(17):2947–2953, 1999.
- [8] H Apfelbaum, D Apfelbaum, R Woods, and Eli Peli. 41.2: The effect of edge filtering on vision multiplexing. In *SID Symposium Digest of Technical Papers*, volume 36, pages 1398–1401. Wiley Online Library, 2005.
- [9] H. Apfelbaum, D. Apfelbaum, Russell L Woods, and Eli Peli. 41.2: The effect of edge filtering on vision multiplexing. *SID Symposium Digest of Technical Papers*, 36(1):1398–1401, 2005.
- [10] Henry L. Apfelbaum, Doris H. Apfelbaum, Russell L. Woods, and Eli Peli. The effect of edge filtering on inattention blindness. *J. Vis.*, 5(8):547–547, 9 2005.
- [11] Henry L. Apfelbaum, Doris H. Apfelbaum, Russell L. Woods, and Eli Peli. Inattention blindness and augmented vision displays: Effects of cartoon like filtering and attended scene. ophthalmic physiol optics. Technical report, Harvard Medical School, 2007.

- [12] Rétine artificielle et lunettes à réalité augmentée au programme du colloque Optic 2000. Rétine artificielle et lunettes à réalité augmentée au programme du colloque optic 2000. Acuité, october 2009.
- [13] Jean Babaud, Andrew P Witkin, Michel Baudin, and Richard O Duda. Uniqueness of the gaussian kernel for scale-space filtering. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (1):26–33, 1986.
- [14] Nicolas Baltenneck. *Se mouvoir sans voir. Incidences de l’environnement urbain sur la perception, la représentation mentale et le stress lors du déplacement de la personne aveugle*. PhD thesis, Université Lumière-Lyon II, 2010.
- [15] James Barabas, Russell L Woods, Robert B Goldstein, and Eli Peli. Perception of collisions while walking in a virtual environment with simulated peripheral vision loss. *Journal of Vision*, 4(8):806–806, 2004.
- [16] Eric Beaubert, Franck Pariguet, and Stéphane Taboulot. *Manuel de l’opticien*. ISBN 2-224-02813-X. Editions Maloine, 2005.
- [17] Franck Bequet. La dégénérescence maculaire liée à l’age. In Vision Presse, editor, *3D Vision*, number 2, pages 18–21, 2005.
- [18] Jean Baptiste Bernard. *Psychophysique de la lecture en basse vision: Simulation de scotome, modelisation et developpement d’aides visuelles*. PhD thesis, Université de la Méditerranée Aix-Marseille II, fevrier 2009.
- [19] Jean Baptiste BERNARD, Fermin MOSCOSO DEL PRADO MARTIN, Anna MONTAGNINI, and Eric CASTET. A model of optimal oculomotor strategies in reading for normal and damaged visual fields. In *Deuxième conférence française de Neurosciences Computationnelles, Neurocomp08*, Marseille France, 2008.
- [20] Jean Baptiste Bernard, Anne Catherine Scherlen, and Eric Castet. Page mode reading with simulated scotomas a modest effect of interline spacing on reading speed. *vision research*, 47:3447 3459, 2007.
- [21] Jean-Baptiste Bernard, Anne-Catherine Scherlen, Françoise Vitu-Thibault, and Eric Castet. Effect of line spacing on reading speed in normally-sighted subjects with an artificial scotoma. *Journal of VISION*, 7(9):518–518, 2007.
- [22] Jean Baptiste Bernard, Emilien Tlapale, Geraldine Faure, Eric Castet, and Pierre Kornprobst. Navisio: Towards an integrated reading aid system for low vision patients. In

Workshop on Computer Vision Applications for the Visually Impaired, Marseille France, 2008.

[23] Pascal Bertolino. *Contribution des pyramides irrégulières en segmentation d'images multirésolution*. PhD thesis, Institut National Polytechnique de Grenoble-INPG, 1995.

[24] I Boaventure and Adilson Gonzaga. Method to evaluate the performance of edge detector. In *Proceedings of the The Brazilian Symposium on Computer Graphics and Image Processing*, 2009.

[25] Nicolas Bonnier, Francis Schmitt, Michael Hull, and Christophe Leynadier. Spatial and color adaptive gamut mapping: A mathematical framework and two new algorithms. In *Color and Imaging Conference*, volume 2007, pages 267–272. Society for Imaging Science and Technology, 2007.

[26] Muriel Boucart, Jean-François Dinon, Pascal Despretz, Thomas Desmettre, Katrine Hladiuk, Aude Oliva, et al. Recognition of facial emotion in low vision: a flexible usage of facial features. *Visual neuroscience*, 25(4):603–9, 2008.

[27] E. Bourennane, P. Gouton, M. Paindavoine, and F. Truchetet. Generalization of canny deriche for detection of noisy exponential edge. *Signal processing*, 82:1317–1328, 2002.

[28] Alex R Bowers, Gang Luo, Noa M Rensing, and Eli Peli. Evaluation of a prototype minified augmented-view device for patients with impaired night vision*. *Ophthalmic and Physiological Optics*, 24(4):296–312, 2004.

[29] AR Bowers, D Apfelbaum, and E Peli. Driving habits of visually impaired drivers who use bioptic telescopes. *Investigative Ophthalmology and Visual Science*, 45(5):4585, 2004.

[30] AR Bowers, GANG Luo, and E Peli. Functionally relevant illumination levels for evaluation of a new night vision device. *Invest. Ophthalmol. Vis. Sci*, 44:2772, 2003.

[31] Erin Brady, Meredith Ringel Morris, Yu Zhong, Samuel White, and Jeffrey P Bigham. Visual challenges in the everyday lives of blind people. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2117–2126. ACM, 2013.

[32] D.K. Lee L. Itti C. Koch J. Braun. Attention activates winner take all competition among visual filters. *Nature Neuroscience*, 2:375 381, 1999.

[33] Gustav J Braun and Mark D Fairchild. General-purpose gamut-mapping algorithms: evaluation of contrast-preserving rescaling functions for color gamut mapping. In *Color and Imaging Conference*, volume 1999, pages 167–172. Society for Imaging Science and Technology, 1999.

- [34] Peter Burt and Edward Adelson. The laplacian pyramid as a compact image code. *Communications, IEEE Transactions on*, 31(4):532–540, 1983.
- [35] Ozan Cakmakci and Jannick Rolland. Head-worn displays: a review. *Journal of Display Technology*, 2(3):199–216, 2006.
- [36] Aurélie Calabrèse, Jean-Baptiste Bernard, Louis Hoffart, Géraldine Faure, Fatiha Barouch, John Conrath, and Eric Castet. Small effect of interline spacing on maximal reading speed in low-vision patients with central field loss irrespective of scotoma size. *Investigative Ophthalmology & Visual Science*, 51(2):1247–1254, fevrier 2010. Association for Research in Vision and Ophthalmology.
- [37] John Canny. A computational approach to edge detection. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 8(6):679–698, 1986.
- [38] John Francis Canny. *Finding edges and lines in images*. PhD thesis, MIT Artificial Intelligence Laboratory, 1983.
- [39] Eric Castet and Aurélie Calabrese. psychophysique de la lecture chez les patients basse vision porteurs de scotomes maculaire http://ifrath.fr/wp-content/uploads/2009/01/calabrese_20090121.pdf.
- [40] Sébastien Chabrier, Hélène Laurent, Christophe Rosenberger, and Bruno Emile. Comparative study of contour detection evaluation criteria based on dissimilarity measures. *Journal on Image and Video Processing*, 2008:2, 2008.
- [41] BHABATOSH CHANDA, MALAY K. KUNDU, and Y. VANI PADMAJA. A multi-scale morphologic edge detector. *pattern recognition*, 31(10):1469–1478, 1998.
- [42] Heng-Da Cheng, XH Jiang, Ying Sun, and Jingli Wang. Color image segmentation: advances and prospects. *Pattern recognition*, 34(12):2259–2281, 2001.
- [43] Ming-Ming Cheng, Guo-Xin Zhang, Niloy J Mitra, Xiaolei Huang, and Shi-Min Hu. Global contrast based salient region detection. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 409–416. IEEE, 2011.
- [44] N Congdon, B O’colmain, CC Klaver, R Klein, B Munoz, DS Friedman, J Kempen, HR Taylor, and P Mitchell. Causes and prevalence of visual impairment among adults in the united states. *Archives of ophthalmology*, 122(4):477–485, 2004.
- [45] Christain Corbé. La forte myopie. In Vision Presse, editor, *3D Vision*, number 5, pages 4 – 6, January 2007.

- [46] Christine A Curcio, Kenneth R Sloan, Robert E Kalina, and Anita E Hendrickson. Human photoreceptor topography. *Journal of Comparative Neurology*, 292(4):497–523, 1990.
- [47] Éric Dinet and Anthony Bartholin. A biologically plausible model for selective color image processing. In *Proceedings of the Tenth IASTED International Conference on Computer Graphics and Imaging*, pages 184–189. ACTA Press, 2008.
- [48] Eric Dinet and Emmanuel Kubicki. A selective attention model for predicting visual attractors. In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pages 697–700. IEEE, 2008.
- [49] Lijun Ding and Ardeshir Goshtasby. On the canny edge detector. *Pattern Recognition*, 34(3):721–725, 2001.
- [50] David Drascic and Paul Milgram. Perceptual issues in augmented reality. In *Electronic Imaging: Science & Technology*, pages 123–134. International Society for Optics and Photonics, 1996.
- [51] A. T. Duchowski and T. D. Eaddy. A gaze-contingent display compensating for scotomata. In *EUROGRAPHICS*, 2009.
- [52] Andrew T. Duchowski. *eye tracking methodology theory and practice*. springer, 2007.
- [53] ANDREW T. DUCHOWSKI and ARZU COLTEKIN. Foveated gaze contingent displays for peripheral lod management, 3d visualization, and stereo imaging. *ACM Transactions on Multimedia Computing, Communications and Applications*, 3(4):24, decembre 2007.
- [54] Sashah Eftekhari. *Augmented reality and scene examination*. PhD thesis, University of Birmingham, 2011.
- [55] M.R. Everingham, B.T. Thomas, and T. Troscianko. Head-mounted mobility aid for low vision using scene classification techniques. *THE INTERNATIONAL JOURNAL OF VIRTUAL REALITY*, 3(3):0, 1998.
- [56] Pedro Felzenszwalb, David McAllester, and Deva Ramanan. A discriminatively trained, multiscale, deformable part model. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.
- [57] Elisabeth M. Fine and Eli Peli. Enhancement of text for the visually impaired. *J. OPT. SOC. AM. A*, 12(7):1439 1447, 1995.

- [58] Timothy E. Fisher. programmable remapper with single flow architecture us patent: Us 5208872, 1993.
- [59] E. Fontaine, Guillaume Tatur, and Edvvige Pissaloux. Experiments in virtual navigation as a step in the development of a navigation tool for blind people. In *Assistive Technology for All Ages. Conference & Workshop on Assistive Technologies for People with Vision & Hearing Impairments*, 2007.
- [60] P Forssen and Erik Ringaby. Rectifying rolling shutter video from hand-held devices. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 507–514. IEEE, 2010.
- [61] Per-Erik Forssén, Erik Ringaby, and Johan Hedborg. Tutorial computer vision on rolling shutter cameras cvpr 2012 <http://www.cvl.isy.liu.se/education/tutorials/rolling-shutter-tutorial>.
- [62] Société française d’ophtalmologie and Jean-Francois Risse. *Exploration de la fonction visuelle: applications au domaine sensoriel de l’oeil normal et en pathologie*. Masson, 1999.
- [63] Matthew Fullerton and Eli Peli. P-35: Distinguished poster paper: Mpeg-based image enhancement for the visually impaired: Implementation on a general-purpose pc platform. In *SID Symposium Digest of Technical Papers*, volume 36, pages 402–405. Wiley Online Library, 2005.
- [64] Davide Gadia, Cristian Bonanomi, Maurizio Rossi, Alessandro Rizzi, and Daniele Marini. Color management and color perception issues in a virtual reality theater. In *Electronic Imaging 2008*, pages 68030S–68030S. International Society for Optics and Photonics, 2008.
- [65] M. A. García-Pérez and Eli Peli. Simple non-invasive measurement of rapid eye vibration. *Journal of Sound and Vibration*, 262(4):877–888, 2003.
- [66] M García-Pérez and Eli Peli. Motion perception under involuntary eye vibration. *Perception*, 31:S177, 2002.
- [67] Wilson S. Geisler and Jiri Najemnik. Eye movement statistics in humans are consistent with an optimal search strategy. *Journal of Vision*, 8(3):1–14, 2008.
- [68] Th Gevers and AWM Smeulders. Edge steered region segmentation by photometric color invariant. pages 869–877, 1997.
- [69] Theo Gevers. Reflectance-based classification of color edges. *Computer Vision, IEEE International Conference on*, 2:856, 2003.

- [70] Robert G Giorgi, Russell L Woods, Robert B Goldstein, Eli Peli, Tamara Bond, and Randy D Easton. Extended wearing trial of peripheral prism device for homonymous hemianopia. 2004.
- [71] Catherine Goillot and Pierre Mormiche. Les enquêtes handicaps-incapacités-dépendances de 1998 et 1999. Institut national de la statistique et des études économiques INSEE, october 2003.
- [72] RB Goldstein, E Peli, H Apfelbaum, and R Hier. Dynamic control of magnified image for low vision observers. In *2002 Annual Meeting Abstract and Program Planner [on CD-ROM]. Association for Vision and Ophthalmology. Presentation*, volume 3803, 2002.
- [73] Robert B Goldstein, Henry Apfelbaum, Gang Luo, and Eli Peli. 37.3: Dynamic magnification of video for people with visual impairment. In *SID Symposium Digest of Technical Papers*, volume 34, pages 1152–1155. Wiley Online Library, 2003.
- [74] Robert B. Goldstein, Henry Apfelbaum, Russell L. Woods, Robert Giorgi, Eli Peli, James Barabas Robert, and James Barabas. Tracking the line of primary gaze in a walking simulator: Modeling and calibration, 2004.
- [75] Robert B. Goldstein, Eli Peli, Shabtai Lerner, and Gang Luo. Eye movements while watching video: comparisons across viewer groups. *J. Vis.*, 4(8):643–643, 8 2004.
- [76] Javier-Flavio Viguera Gomez, Gilles Simon, and Marie-Odile Berger. Calibration errors in augmented reality: A practical study. In *Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 154–163. IEEE Computer Society, 2005.
- [77] Daniel D Gottlieb. Method of using a prism in lens for the treatment of visual field loss us patent 4,779,972, 1988.
- [78] Carine Grasset-Simon. Définition et étude des pyramides généralisées nd: application pour la segmentation multi-échelle d’images 3d. *These de doctorat, Université de Poitiers*, 2006.
- [79] ROBERT HARPER, LOUISE CULHAM, and CHRISTINE DICKINSON. head mounted video magnification devices for low vision rehabilitation : a comparison with existing technology. *Br Journal Ophthalmology*, 83:495 500, 1999.
- [80] Abdelsalam Helal, Steven Edwin Moore, and Balaji Ramachandran. Drishti: An integrated navigation system for visually impaired and disabled. In *Wearable Computers, 2001. Proceedings. Fifth International Symposium on*, pages 149–156. IEEE, 2001.

- [81] Stephen HICKS. An optical device for the visually impaired, August 31 2012. WO Patent 2,012,114,123.
- [82] Richard L Holloway. Registration error analysis for augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4):413–432, 1997.
- [83] Jun Wei Hsieh, Ming Tat Ko, Hong Yuan Mark Liao, and Kuo Chin Fan. a new wavlet based edge detector via constrained optimization. *image and vision computing*, 15:511–527, 1997.
- [84] Laurent Itti. accueil ilab: <http://ilab.usc.edu/>.
- [85] Laurent Itti. ilab bibliotheque: <http://ilab.usc.edu/publications/l.itti.html>.
- [86] Laurent Itti. Ilab recherche <http://ilab.usc.edu/research/>.
- [87] Laurent Itti. Models of bottom-up attention and saliency. *Neurobiology of Attention*, Elsevier, 2005.
- [88] Laurent ITTI and Christof KOCH. a comparison of feature combination stratégies for saliency based visual attention systems. *Society of Photo-Optical Instrumentation Engineers, Bellingham, WA, INTERNATIONAL*, 0:0, 1998. SPIE, Bellingham WA, ETATS-UNIS (1999) (Monographie).
- [89] Jean Michel Jolion. analyse d'image: le modèle pyramidal. *traitement du signal*, 7(1):0, 1990.
- [90] J Adam Jones, Evan A Suma, David M Krum, and Mark Bolas. Comparability of narrow and wide field-of-view head-mounted displays for medium-field distance judgments. In *Proceedings of the ACM Symposium on Applied Perception*, pages 119–119. ACM, 2012.
- [91] Mamta Juneja and Parvinder Singh Sandhu. Performance evaluation of edge detection techniques for images in spatial domain. *methodology*, 1(5):614–621, 2009.
- [92] Roland Kalberer. *Lunettes à ultrasons pour malvoyants*. PhD thesis, éceole polytechnique federale de Lausanne, 1990.
- [93] Jeonghoon Kim and Eli Peli. Mpeg-based image enhancement for the visually impaired. *Optical engineering*, 43(6):1318 1328, 2003.
- [94] H Konik, B Laget, and M Calonnier. Segmentation d'images par utilisation de pyramides à bases locales. *Traitement du Signal*, 10(4):283–296, 1993.
- [95] Caroline Kovarski. *L'opticien-lunetier guide théorique et pratique*. ISBN 2-7430-0651-X. Edition Lavoisier, 2005.

- [96] Gordon E Legge, Gary S Rubin, Denis G Pelli, and Mary M Schleske. Psychophysics of reading—ii. low vision. *Vision research*, 25(2):253–265, 1985.
- [97] L Li, E Peli, and WH Warren. Heading perception of tunnel-vision patients. In *INVESTIGATIVE OPHTHALMOLOGY & VISUAL SCIENCE*, volume 41, pages S797–S797. ASSOC RESEARCH VISION OPHTHALMOLOGY INC 9650 ROCKVILLE PIKE, BETHESDA, MD 20814-3998 USA, 2000.
- [98] Wu Li and Charles D. Gilbert. Global contour saliency and local colinear interactions. *Journal of Neurophysiology*, 88:2846–2856, 2002.
- [99] Zile Liu and Rachid Hamdani. Procédés et dispositif pour la création d’images rétiniennes utilisant le stigmatisme de deux foyers d’un dioptre sensiblement elliptique brevet fr 2 866 123 - a1, 2004.
- [100] Zile Liu and Rachid Hamdani. Method and device for generating retinal image using the stigmatism of two foci of substantially elliptical sight brevet w0 2005/078510 a2, 2005.
- [101] Zile Liu and Rachid Hamdani. Dispositif portable d’affichage tête haute et de réalité augmentée brevet fr 2 941 786 - a1, 2008.
- [102] Zile Liu and Rachid Hamdani. Portable augmented reality head up display device brevet wo 2010/089495 a2, 2010.
- [103] Mark A Livingston, Joseph L Gabbard, J Edward Swan II, Ciara M Sibley, and Jane H Barrow. Basic perception in head-worn augmented reality displays. In *Human Factors in Augmented Reality Environments*, pages 35–65. Springer, 2013.
- [104] Carlos Lopez-Molina, Bernard De Baets, Humberto Bustince, Josean Sanz, and Ederne Barrenechea. Multiscale edge detection based on gaussian smoothing and edge tracking. *Knowledge-Based Systems*, 2013.
- [105] David G Lowe. Organization of smooth image curves at multiple scales. *International Journal of Computer Vision*, 3(2):119–130, 1989.
- [106] Zhong-Lin Lu, Luis A. Lesmes, and George Sperling. Perceptual motion standstill in rapidly moving chromatic displays. *proceedings of national academy of sciences of united states of america*, 96(26):15734–15739, 1999.
- [107] Gang Luo and Eli Peli. Use of an augmented-vision device for visual search by patients with tunnel vision. In *Investigative Ophthalmology & Visual Science*, page 4152–4159, 2006.

- [108] Gang Luo, Noa Rensing, Evan Weststrate, and Eli Peli. Registration of an on axis see-through head mounted display and camera system. *Optical Engineering*, 2005.
- [109] Gang Luo, Russell L. Woods, and Eli Peli. Collision judgment when using an augmented vision head mounted display device, 2009.
- [110] Michael Randolph Maire. *Contour Detection and Image Segmentation*. PhD thesis, Electrical Engineering and Computer Sciences University of California at Berkeley, 2009.
- [111] Sharada Mani, G. SyamPrasad, O.Rama Devi, and C.Naga Raju. A novel edge detection technique for color images. *International Journal of Advanced Engineering & Application*, 0:235–238, 2010.
- [112] David MARR. *Vision a computational investigation into the human representation and processing of visual information*. The MIT Press, 1982.
- [113] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proc. 8th Int’l Conf. Computer Vision*, volume 2, pages 416–423, July 2001.
- [114] Émilien Tlapale, Jean-Baptiste Bernard, Éric Castet, and Pierre Kornprobst. The solaire project: A gaze-contingent system to facilitate reading for patients with scotomas. Technical Report 0326, INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE, octobre 2006.
- [115] Kirk Moffitt. Head-mounted display image configurations, document non publié disponible à http://www.kirkmoffitt.com/hmd_image_configurations.pdf juin 2012, 2008.
- [116] Farzin Mokhtarian and Riku Suomela. Robust image corner detection through curvature scale space. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 20(12):1376–1381, 1998.
- [117] Jan Morovic and M Ronnier Luo. Calculating medium and image gamut boundaries for gamut mapping. *Color research and application*, 25(6):394–401, 2000.
- [118] Ehsan Nadernejad, Sara Sharifzadeh, and Hamid Hassanpour. Edge detection techniques: evaluations and comparisons. *Applied Mathematical Sciences*, 2(31):1507–1520, 2008.
- [119] Hong Shan Neoh and Asher Hazanchuk. Adaptive edge detection for real-time video processing using fpgas. *Global Signal Processing*, 2004.

- [120] Alan F Newell and Peter Gregor. “user sensitive inclusive design”—in search of a new paradigm. In *Proceedings on the 2000 conference on Universal Usability*, pages 39–44. ACM, 2000.
- [121] Jonathan Nissanov. Method and apparatus for central field loss mapping and optimization of image presentation based upon mapped parameters brevet wo96 34555, 1996.
- [122] N.Mandran, B.Meillon, and A.Pellegrin, editors. *oculométrie*, Paris France, 2007.
- [123] Alex K Nugent, Rajesh N Keswani, Russell L Woods, and Eli Peli. Contour integration in peripheral vision reduces gradually with eccentricity. *Vision research*, 43(23):2427–2437, 2003.
- [124] Alex K Nugent, Russell L Woods, and Eli Peli. Flanker size affects visual lateral interactions. In *Association of Research Vision Ophthalmology 2002 annual meeting, CD-ROM*, volume 4718, 2002.
- [125] World Health Organization. the use of residual vision by visually disabled person, euro reports and studies 41. Regional Office for Europe WHO, January 1981. Copenhagen, Denmark.
- [126] World Health Organization. Global data on visual impairments, 2012. WHO/NMH/PBD/12.01.
- [127] Mohammadreza Asghari Oskoei and Huosheng Hu. A survey on edge detection methods. Technical report, Technical Report CES, 2010.
- [128] Charles B Owen, Ji Zhou, Arthur Tang, and Fan Xiao. Display-relative calibration for optical see-through head-mounted displays. In *Mixed and Augmented Reality, 2004. ISMAR 2004. Third IEEE and ACM International Symposium on*, pages 70–78. IEEE, 2004.
- [129] Giuseppe Papari, Patrizio Campisi, Nicolai Petkov, and Alessandro Neri. A biologically motivated multiresolution approach to contour detection. *EURASIP Journal on Applied Signal Processing*, 2007(1):119–119, 2007.
- [130] J.R. Parker. *ALGORITHMS FOR IMAGE PROCESSING AND COMPUTER VISION Advanced edge detection techniques the canny and Shen Castan methods*. WILEY COMPUTER PUBLISHING, 2010.
- [131] Derrick Parkhurst, Eugenio Culurciello, and Ernst Niebur. Evaluating variable resolution displays with visual search: Task performance and eye movements. In *Proceedings of the 2000 symposium on Eye tracking research & applications*, pages 105–109. ACM, 2000.

- [132] Derrick Parkhurst, Klinto Law, and Ernst Niebur. Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, 42:107–123, 2002.
- [133] D Pascolini, SP Mariotti, GP Pokharel, R Pararajasegaram, D Etya, A-D Négrel, and S Resnikoff. 2002 global update of available data on visual impairment: a compilation of population-based prevalence studies. *Neuro-Ophthalmology*, 11(2):67–115, 2004.
- [134] Donatella Pascolini and Silvio Paolo Mariotti. Global estimates of visual impairment: 2010. *British Journal of Ophthalmology*, 96(5):614–618, 2012.
- [135] E Peli, RB Goldstein, RL Woods, JH Kim, and Y Yitzhaky. Wide band enhancement of tv images for the visually impaired. *Invest Ophthalmol Vis Sci*, 45, 2004.
- [136] Eli Peli. Vision rehabilitation laboratory, harvard, eli peli <http://www.eri.harvard.edu/faculty/peli/>.
- [137] Eli Peli. Visual issues in the use of a head-mounted monocular display. *Optical Engineering*, 29(8):883–892, 1990.
- [138] Eli Peli. *Vision models for target detection and recognition*, volume 2. World scientific, 1995.
- [139] Eli Peli. The visual effects of head-mounted display (hmd) are not distinguishable from those of desk-top computer display. *Vision Research*, 38(13):2053–2066, 1998.
- [140] Eli Peli. Augmented vision for central scotoma and peripheral field loss. *Vision Rehabilitation: Assessment, Intervention and Outcomes*. Lisse, Netherlands: Swets & Zeitlinger, pages 70–4, 2000.
- [141] Eli Peli. Vision multiplexing: an engineering approach to vision rehabilitation device development. *Optometry & Vision Science*, 78(5):304–315, 2001.
- [142] Eli Peli. wide band image enhancement eli peli us patent 6,611,618, 2003.
- [143] Eli Peli, editor. *Extended wearing trial of trifield prism visual aid for tunnel vision among patients with retinitis pigmentosa or choroideremia*, 2004.
- [144] Eli Peli. wide band image enhancement eli peli us patent 7,280,704, 2007.
- [145] Eli Peli, Alex R Bowers, Aaron J Mandel, K Higgins, Rise B Goldstein, and L Bobrow. Design for simulator performance evaluations of driving with vision impairments and visual aids. *Transportation Research Record: Journal of the Transportation Research Board*, 1937(1):128–135, 2005.

- [146] Eli Peli, Gang Luo, Alex Bowers, and Noa Rensing. Applications of augmented vision head mounted systems in vision rehabilitation. *Society for Information Display*, 15(12):1037 1045, 2007.
- [147] Eli Peli, Gang Luo, Alex Bowers, and Noa Rensing. Development and evaluation of vision multiplexing devices for vision impairments. *International Journal on Artificial Intelligence Tools*, 18(03):365–378, 2009.
- [148] Eli Peli, Gang Luo, and Lee Lichtenstein. Collision judgment when viewing minified images through a hmd visual field expander. *Investigative Ophthalmology & Visual Science*, 50(9):4509 4515, septembre 2009.
- [149] Eli Peli and Fernando Vargas Martin. Augmented-view for restricted visual field: Multiple device implementations. *Optometry & Vision Science*, 11:715 723, 2002.
- [150] Eli Peli and Fernando Vargas Martin. Eye movements of patients with tunnel vision while walking. *Investigative Ophthalmology & Visual Science*, 47(12):5295 5302, decembre 2006.
- [151] Eli Peli and Russell L. Woods. Image enhancement for impaired vision: the challenge of evaluation. *International journal of artificial intelligence tools architectures languages algorithms*, 18(3):415–438, 2009.
- [152] Robert J. Peters, Asha Iyer, Laurent Itti, and Christof Koch. Components of bottom up gaze allocation in natural images. *Vision Research*, 45:2397 2416, 2005.
- [153] Edwige Pissaloux, Hichem Bouayed, and Samer Abdallah. Comparison of some vision systems for blinds. In *IAPR Workshop on Machine Vision Applications*. The University of Tokyo.Japan, novembre 2000.
- [154] Fatih Porikli. Integral histogram: A fast way to extract histograms in cartesian spaces. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 829–836. IEEE, 2005.
- [155] projet Lasdop. *Projet lasdop*, Nimes, 0000.
- [156] Graham Pullin. *Design meets disability*. MIT press, 2009.
- [157] Warren Robinett and Richard Holloway. The visual display transformation for virtual reality. *Presence: Teleoperators and Virtual Environments*, 4(1):1–23, 1995.
- [158] Jannick P Rolland and Ozan Cakmakci. The past, present, and future of head-mounted display designs. In *Photonics Asia 2004*, pages 368–377. International Society for Optics and Photonics, 2005.

- [159] Jannick P Rolland and Terry Hopkins. A method of computational correction for optical distortion in head-mounted displays. *Univ. N. Carolina at Chapel Hill Tech Rpt. 93*, 45, 1993.
- [160] JP Rolland and Hong Hua. Head-mounted display systems. *Encyclopedia of optical engineering*, pages 1–13, 2005.
- [161] Ramón Román-Roldán, Juan Francisco Gómez-Lopera, Chakir Atae-Allah, José Martínez-Aroza, and Pedro Luis Luque-Escamilla. A measure of quality for evaluating methods of segmentation and edge detection. *Pattern recognition*, 34(5):969–980, 2001.
- [162] NC Ross, AR Bowers, and E Peli. Consideration of optical scotomas in designing visual field expansion devices. *Investigative Ophthalmology and Visual Science*, 50(5):4734, 2009.
- [163] I.A. Rybak, V.I. Gusakova, A.V. Golovan, L.N. Podladchikova, and N.A. Shevtsova. A model of attention guided visual perception and recognition. *Vision Research*, 38:2387–2400, 1998.
- [164] Richard D. Juday Jeffrey B. Sampsell. programmable remapper for image processing richard d. juday jeffrey b. sampsell us patent 5,067,019, 1991.
- [165] Marie-Sylvie Sander. La population en situation de handicap visuel en france. Observatoire Régional de la Santé des pays de la Loire, Juillet 2005.
- [166] A. Scherlen and V. Gautier. Improvement of reading performance with a visual scotoma by synchronizing eye movements to visual information availability. *Investigative Ophthalmology & Visual Science*, 45:0, 2004.
- [167] Anna-Catherine Scherlen and Vincent Gauthier. procédé de démasquage de l’information visuelle présente sous une zone peu ou pas fonctionnelle et le dispositif de mise en oeuvre, 2004.
- [168] Anne Catherine Scherlen, Jean-Baptiste Bernard, Aurélie Calabrese, and Eric Castet. Page mode reading with simulated scotomas: Oculo-motor patterns. *Vision Research*, 48:1870–1878, 2008.
- [169] Anne Catherine Scherlen and Vincent Gautier. The difference between mean saccade size and scotoma size is a predictive factor for reading speed. In *International Congress Series*, number 1282, page 704–708, 2005.
- [170] Anne-Catherine Scherlen and Vincent Gautier, editors. *A new concept for visual aids ViSAR Visual Signal Adaptive Restitution*, number 27 in Engineering in Medicine and

Biology, Shanghai, China, septembre 2005. IEEE. Engineering in Medicine and Biology 27th Annual Conference Shanghai, China, September 1-4, 2005.

[171] Susanne Schuett, RobertW. Kentridge, Josef Zihl, and Charles A. Heywood. Adaptation of eye movements to simulated hemianopia in reading and visual exploration: Transfer or specificity? *Neuropsychologia*, 47:1712–1720, 2009.

[172] Susanne Schuett, RobertW. Kentridge, Josef Zihl, and Charles A. Heywood. Are hemianopic reading and visual exploration impairments visually elicited? new insights from eye movements in simulated hemianopia. *Neuropsychologia*, 47:733–746, 2009.

[173] Kristen Shinohara and Jacob O Wobbrock. In the shadow of misperception: assistive technology use and social interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 705–714. ACM, 2011.

[174] Dave Stampe. Heuristic filtering and reliable calibration methods for video based pupil tracking systems. *Behavior Research Methods*, 25:137–142, 1993. 10.3758/BF03204486.

[175] Kent A. Stevens. method and apparatus for measuring and correcting metamorphosis. kent a. stevens us patent us 5892570, 1999.

[176] Brian Douglas Strege. *Using Tiled Head-Mounted Displays with a Single Video Source*. PhD thesis, Citeseer, 2009.

[177] Baris Sumengen and BS Manjunath. Multi-scale edge detection and image segmentation. In *European signal processing conference (EUSIPCO)*. Citeseer, 2005.

[178] Richard Szeliski. *Computer vision: algorithms and applications*. Springer, 2011.

[179] Guillaume TATUR. *Conception d'un système de vision par phosphènes*. PhD thesis, Académie de Montpellier, université Montpellier II Sciences et techniques du Languedoc, 2011.

[180] Kaveri A Thakoor, Sophie Marat, Patrick J Nasiatka, Ben P McIntosh, Furkan E Sahin, Armand R Tanguay, James D Weiland, and Laurent Itti. Attention biased speeded up robust features (ab-surf): A neurally-inspired object recognition algorithm for a wearable aid for the visually-impaired. In *Multimedia and Expo Workshops (ICMEW), 2013 IEEE International Conference on*, pages 1–6. IEEE, 2013.

[181] A Thuewissen. Image sensor architectures for digital cinematography. Technical report, Technical report, DALSA Digital Cinema, 2005.

[182] George T Timberlake, Martin A Mainster, Eli Peli, Reed A Augliere, Edward A Essock, and Lawrence E Arend. Reading with a macular scotoma. i. retinal location of

scotoma and fixation area. *Investigative Ophthalmology & Visual Science*, 27(7):1137–1147, 1986.

[183] Anne M. Treisman and Garry Gelade. *A feature integration theory of attention*, volume 1. Cognitive psychology, 1980.

[184] FAA USDOT. Human factors design standard (hfds). *Rep. DOT/FAA/CT-03/05 HF-STD-001*, 2003.

[185] Joost Van De Weijer, Theo Gevers, and Arnold WM Smeulders. Robust photometric invariant features from the color tensor. *Image Processing, IEEE Transactions on*, 15(1):118–127, 2006.

[186] F Vargas-Martin and E Peli. Eye movements patterns in walking hemianopic patients. *Investigative Ophthalmology and Visual Science*, 43(12):3809, 2002.

[187] Fernando Vargas-martín and Eli Peli. P-16: Augmented view for tunnel vision: Device testing by patients in real environments, 2001.

[188] J. van de Weijer, Th. Gevers, and J. M. Geusebroek. Color edge detection by photometric quasi-invariants. In *Proceedings of the Ninth IEEE International Conference on Computer Vision - Volume 2*, ICCV '03, pages 1520–, Washington, DC, USA, 2003. IEEE Computer Society.

[189] Jeremy M. Wolfe. Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin Review*, 1(2):202–238, 1994.

[190] wolfgang Einhauser, Nathan Mundhenk, Pierre Baldi, Chirstof Koch, and Laurent Itti. A bottom up model of spatial attention predicts human error patterns in rapid scene recognition. *journal of vision*, 7(10):1–13, 2007.

[191] Bonita P. H. Wong, Russell L Woods, and Eli Peli. Stereoacuity at distance and near. *Optometry & Vision Science*, 79(12):771–778, 2002.

[192] RL Woods and E Peli. Development of a novel optical aid for people with severely restricted visual fields. *Investigative Ophthalmology and Visual Science*, 43(12):3799, 2002.

[193] RL Woods, JC Shieh, L Bobrow, A Vora, J Barabas, RB Goldstein, and E Peli. Perceived collision with an obstacle in a virtual environment. *Investigative Ophthalmology and Visual Science*, 44(5):4321, 2003.

[194] Russell Woods, Robert Goldstein, and Daniel Stringer. Development and testing of trifold glasses for people with severely restricted visual fields.: Poster# 158. *Optometry & Vision Science*, 79(12):187, 2002.

- [195] Russell L Woods, Ivonne Fetchenheuer, Fernando Vargas-martín, and Eli Peli. The impact of non-immersive head mounted displays (hmds) on the visual field. *J Soc Inform Display*, 2003:0, 2003.
- [196] Russell L Woods, Aaron J Mandel, James Barabas, Robert B Goldstein, and Eli Peli. Making virtual reality more real and the perception of potential collisions. *Journal of Vision*, 4(8):814, 2004.
- [197] Russell L Woods and PremNandhini Satgunam. Television, computer and portable display device use by people with central vision impairment. *Ophthalmic and Physiological Optics*, 31(3):258–274, 2011.
- [198] Alfred L. Yarbus. *eye movements and vision*. Plenum press new york, 1967.
- [199] Bing Zeng. Reduction of blocking effect in dct coded images using zero masking techniques. *Signal Processing*, 79(2):205 211, 1999.